THE UNIVERSITY *of* EDINBURGH

Investigation into Early Growth Response 1 in colorectal disease: a study of *EGR1* expression in colorectal tissue and novel protein interactions in cancer cells

Gráinne Mary Gernon

Submitted to The University of Edinburgh

for the degree of

Doctor of Philosophy

School of Medicine and Veterinary Medicine
2012

Declaration

This thesis is presented to the University of Edinburgh for the degree of Doctor of Philosophy.

I, the candidate, declare that this thesis has been composed by myself, that the work contained in this thesis is my own and that the work has not been submitted for any other degree or professional qualification.

Gráinne Gernon

Acknowledgements

This thesis would not have been possible without a great deal of help and guidance from many people. A special thanks to my supervisor Susan for all of her advice, encourgment and support over the course of my PhD. A special thanks also to my second supervisor Elaine for all of her help and guidance. I am very grateful to all of the members of the Colorectal Cancer Genetics Group and Gastrointestinal Unit who aided me in the completion of this thesis, especially Marion, for without her amazing help and endless support in the lab this PhD work could not have been completed. I would also like to thank the members of my thesis committee, Kathyrn and Lesely for their review of my work and guidance.

Thanks is also necessary for everyone on E4, the students and friends at the HGU, who helped solve many a problem, with many a stressfull time relieved by coffee breaks, céilidhs, dinner parties and Friday night drinks in Hectors. A special thanks to Jenny, a great friend, who provided endless quantities of freshly brewed coffee and invaluable 'R' advice, to help keep my sanity! To my amazing flatmates, the Irish girls, who became my Edinburgh family; Aisling, Aisling eile, Sophie and our honoury 'Irish' Anna-Lena, for their endless support, cups of Barrys Tea, rugby matches, dinners, ceoil agus craic! Special thanks to Anna-Lena, in whom I found a constant rock and a life-long friend.

Finally, more thanks that I can express to my parents, for their endless love and support. I could not have done this without them.
To my amazing family; Mum, Dad, Mark, Paul and Courtney, Sinéad, Louise, Sonya, Iain and Conor, Granny, to my aunts, uncles, cousins, neighbours and friends, without whose constant love, support, endless faith and countless candles would have made this PhD and thesis impossible. Thank you for your belief, love and prayers.

Abstract

Introduction: Early growth response 1 (EGR1) is a zinc-finger transcription factor involved in the regulation of cell growth. It can act as either a tumour suppressor or a tumour promoter with a role in the induction of apoptosis in cancer cells by various pathways and is likely to play a role in colorectal cancer (CRC). EGR1 also appears to play a significant role in inflammatory pathways, therefore a possible role in Inflammatory Bowel Disease (IBD) is hypothesised. Patients with IBD have a greater risk of developing CRC, which is increased with duration of symptoms and severity of inflammation and dysplasia. The aim of this study is to determine whether *EGR1* is differentially expressed in diseased colon tissue and to investigate novel EGR1-protein interactions in CRC cell lines.

Methods: The relative *EGR1* expression in CRC cell lines and in normal mucosa and tumours of colorectal cancer patients was determined by qRT-PCR. IBD patient samples were also examined for differential *EGR1* expression levels by qRT-PCR, before and after stimulation with inflammatory mediators. Statistical analysis of the data was performed using 'R' statistical package, with the mixed-model ANOVA. Statistical significance was set at < 0.05. The genotype of three *EGR1* variants was determined in the samples using PCR and sequencing, and the methylation status of regions of the *EGR1* promoter was determined using bisulfite sequencing. A yeast-two hybrid screen was conducted with EGR1 as bait, and screened against a SW480 CRC cell line library. Interesting novel interactions were investigated using immunocytochemistry and immunoprecipitation, as was the novel interaction between EGR1 and NOD2 and between EGR1 and components of the cytoskeleton.

Results: Investigation into the relative *EGR1* mRNA expression in CRC has shown that there is differential expression of *EGR1* between matched normal mucosa and tumour. *EGR1* expression is decreased in IBD patients compared with healthy controls. Induction of *EGR1* by inflammatory stimuli also appears to be aberrant in these patients. The differential expression of *EGR1* was not associated with aberrant methylation of a large region of the *EGR1* promoter in either the CRC or IBD

patients or with the genotype of *EGR1* variants. EGR1 localises to both the cytoplasm and the nucleus in CRC cell lines and this study demonstrate interactions with the IBD susceptibility protein NOD2 and with components of the cyotskeleton. A yeast-two hybrid screen conducted with EGR1 as bait using a CRC cell line library has identified several other novel protein interactions of EGR1 in CRC cell lines.

Conclusion: *EGR1* is differentially expressed in both CRC and IBD, and in the case of IBD shows aberrant activity, suggesting that *EGR1* may play a role in both colorectal diseases. EGR1 interacts with the IBD protein NOD2, and components of the cytoskeleton in CRC cells. Several novel protein interactions with EGR1 have been identified and warrant further study.

Contents

Declaration ii
Acknowledgements iii
Abstract iv
1 Introduction 1
1.1 Colorectal Cancer (CRC) 2
1.1.1 Environmental risk factors for CRC 3
1.1.2 Hereditary forms of CRC 5
1.1.3 Adenoma-carcinoma sequence of CRC 9
1.1.4 Genetic instability in CRC 12
1.1.5 Mutations in CRC 17
1.1.6 Identification of risk alleles in CRC 20
1.2 Inflammatory Bowel Disease (IBD) 22
1.2.1 Environmental risk factors for IBD 24
1.2.2 Genetic risk factors for IBD 27
1.2.3 IBD and CRC 35
1.3 Early Growth Response 1 (EGR1) 37
1.3.1 Structure and promoter of EGR1 38
1.3.2 Induction of EGR1 expression 40
1.3.3 Regulation of EGR1 43
1.3.4 The EGR1 protein and transcription factor 46
1.3.5 EGR1 and cancer 48
1.3.6 EGR1 and inflammation 55
1.4 Preliminary studies of EGR1 58
1.5 Aims and Experimental Design 61
2 Materials and Methods 63
2.1 Biological Material 63
2.1.1 Materials and Solutions 63
2.1.2 Cell culture 64
2.1.3 Treatment of cells 66
2.1.4 Patient Material 66
2.2 RNA protocols 67
2.2.1 RNA extraction 67
2.2.2 Estimation of RNA concentration 67
2.2.3 DNAse I treatment 67
2.2.4 cDNA synthesis 67
2.3 DNA protocols 68
2.3.1 Materials and Solutions 68
2.3.2 DNA extraction 68
2.3.3 Purification of DNA 69
2.3.4 Estimation of DNA concentration 69
2.3.5 Bisulfite treatment of DNA 69
2.4 PCR protocols 70
2.4.1 Oligonucleotides 70
2.4.2 Standard PCR 72

2.4.3 Nested PCR 72
2.4.4 PCR using Bacterial/Yeast plasmids 72
2.4.5 Gel Electrophoresis 72
2.4.6 Quantitative Real-time PCR 73
2.5 Sequence analysis 74
2.5.1 Purification of PCR products 74
2.5.2 DNA sequencing 74
2.5.3 Precipitation of DNA from sequencing reactions 74
2.5.4 Analysis of sequence data 74
2.6 Cloning and Bacterial culture 75
2.6.1 Media and Solutions 75
2.6.2 Plasmids 75
2.6.3 Cloning of *EGR1* into pGEM-T® cloning vector 76
2.6.4 Cloning of *EGR1* into pGBKT7 76
2.6.5 Colony selection and storage 77
2.7 Yeast Culture 78
2.7.1 Media and Solutions 78
2.7.2 Yeast Culture conditions 80
2.7.3 Yeast Transformation 80
2.7.4 Yeast Mating 81
2.7.5 Isolation of plasmids from yeast 81
2.8 Protein Protocols 83
2.8.1 Materials 83
2.8.2 Transfection of cell lines 85
2.8.3 Preparation of protein extracts 85
2.8.4 Western Blot analysis 87
2.8.5 Immunoprecipitation 89
2.9 Immunofluorescence 90
3 Chapter 3 91
3.1 Introduction 91
3.2 Methods 94
3.2.1 Quantitative RT-PCR 94
3.2.2 CRC patients 94
3.2.3 IBD patients 94
3.2.4 Statistical analysis 95
3.2.5 Genotype analysis 95
3.2.6 *KRAS/BRAF* mutation analysis 96
3.2.7 Bisulfite sequencing 96
3.3 Results 98
3.3.1 Relative *EGR1* mRNA expression levels in CRC cell lines and tissue specific expression levels 98
3.3.2 Relative *EGR1* mRNA expression in CRC patients 100
3.3.3 Genotype of *EGR1* variants in matched normal and tumour patient samples 105
3.3.4 *KRAS* and *BRAF* mutation status of CRC patients 109
3.3.5 Methylation status of the *EGR1* promoter 115
3.3.6 Relative *EGR1* mRNA expression levels in IBD patients 117
3.3.7 Genotype analysis of IBD patients for *EGR1* variants 136

3.3.8 Methylation status of the *EGR1* promoter in IBD 139
3.3.9 *NAB2* mRNA expression 141
3.4 Discussion 145
4 Chapter 4 153
4.1 Introduction 153
4.2 Methods 154
4.2.1 Cell Culture 154
4.2.2 siRNA knockdown of EGR1 154
4.2.3 Treatment of cells 155
4.2.4 Preparation of protein extracts 155
4.2.5 Western Blotting 156
4.2.6 Antibodies 157
4.2.7 Immunoprecipitation 158
4.2.8 Immunocytochemistry 159
4.3 Results 161
4.3.1 Determination of accurate tools for EGR1 protein analysis 161
4.3.2 Expression of EGR1 in CRC cells 167
4.3.3 Investigation into a potential interaction of EGR1 and NOD2 173
4.3.4 Investigation into EGR1 interaction with the cytoskeleton 183
4.4 Discussion 187
5 Chapter 5 192
5.1 Introduction 192
5.2 Methods 193
5.2.1 Background of a yeast two-hybrid screen 193
5.2.2 Cloning of *EGR1* into yeast vector pGBKT7 198
5.2.3 Yeast-two hybrid screen 199
5.2.4 Detection of false positives 201
5.3 Results 202
5.3.1 Cloning of EGR1 constructs into the pGBKT7 vector 202
5.3.2 LiAc Yeast Transformation 204
5.3.3 Preliminary Yeast-two hybrid experiments 204
5.3.4 Yeast-two hybrid mating 208
5.3.5 Confirmation of interactions 209
5.3.6 Potential EGR1 interacting proteins 212
5.3.7 Determination of false positives 219
5.3.8 Investigation into the potential interaction of EGR1 and NPM1 220
5.3.9 Investigation into the potential interaction of EGR1 and eEIF1A1 224
5.4 Discussion 228
6 Conclusion 232
7 References 245

Figures

Figure 1.1: Illustration of APC and its protein binding domains. 7
Figure 1.2: Progression of CRC 11
Figure 1.3: Pathways to CRC development 16
Figure 1.4: TGFβ signalling pathway 21
Figure 1.5: Illustration of the IL-23 pathway 29
Figure 1.6: Structure of NOD2 gene and protein 32
Figure 1.7: Development of sporadic CRC versus colitis-associated CRC 36
Figure 1.8: Illustration of EGR1 domains and promoter region 39
Figure 1.9: MAPK pathways involved in *EGR1* activation 42
Figure 1.10: Positive feeback loop with EGR1 and mutant p53 51
Figure 1.11: LPS induction of EGR1 and TF in monocytes 56
Figure 3.1: Illustration of bisulfite sequencing of DNA 97
Figure 3.2: Expression and genotyping of EGR1 in cancer and non-cancer cell lines 99
Figure 3.3: Differential expression of *EGR1* in matched normal mucosa and tumour 101
Figure 3.4: Relative *EGR1* mRNA expression levels in normal mucosa 104
Figure 3.5: Genotype analysis of *EGR1* variants in CRC patients 108
Figure 3.6: *KRAS* and *BRAF* mutation status of CRC patients 110
Figure 3.7: KRAS and BRAF mutation status with relative EGR1 mRNA expression levels in CRC patients. 114
Figure 3.8: Bisulfite sequencing of matched normal and tumour samples 116
Figure 3.9: Relative *EGR1* mRNA expression levels in untreated IBD samples 119
Figure 3.10: Relative *EGR1* mRNA expression levels in CRC cell lines after inflammatory treatment 121
Figure 3.11: Relative *EGR1* mRNA expression levels in healthy controls 123
Figure 3.12: Fold induction of *EGR1* after treatment of healthy control tissue 126
Figure 3.13: Relative *EGR1* mRNA expression levels in ulcerative colitis patients 128
Figure 3.14: Fold induction of *EGR1* after treatment of ulcerative colitis tissue 131
Figure 3.15: Relative *EGR1* mRNA expression levels in Crohn's disease patients 133
Figure 3.16: Fold induction of *EGR1* after treatment of Crohn's disease tissue 135
Figure 3.17: Genotyping of IBD patients and healthy controls 138
Figure 3.18: Bisulfite sequencing of healthy controls and IBD patients 140
Figure 3.19: Relative *NAB2* mRNA expression in CRC cell lines after inflammatory treatment 142
Figure 3.20: Relative *NAB2* mRNA expression levels in IBD and healthy controls 144
Figure 4.1: EGR1 antibody in CRC cell lines 162
Figure 4.2: Knockdown of EGR1 164
Figure 4.3: CRC cells treated with Curcumin and LPS 166
Figure 4.4: Expression of EGR1 in CRC cell lines 168
Figure 4.5: Localisation of EGR1 in CRC cell lines using immunofluorescence 170
Figure 4.6: Localisation of EGR1 in cellular sub-compartments 172
Figure 4.7: Expression of NOD2 in CRC cell lines 174
Figure 4.8: Localisation of NOD2 in CRC cell lines 176

Figure 4.9: Co-localisation of EGR1 and NOD2 in CRC cell lines 178
Figure 4.10: Optimisation of immunoprecipitation 180
Figure 4.11: EGR1 immunoprecipitates with NOD2..... 182
Figure 4.12: EGR1 co-localises with tubulin..... 184
Figure 4.13: EGR1 immunoprecipitates with α- and γ-tubulin 186
Figure 5.1: Illustration of GAL4 mediated activation of reporter genes in yeast two-hybrid system 194
Figure 5.2: Map of the vectors pGBKT7 and pGADT7 197
Figure 5.3: EGR1Fl and EGR1ΔAct constructs..... 203
Figure 5.4: Growth curve of Y187, pGBKT7 and the EGR1 constructs 205
Figure 5.5: Auto-activation assay 207
Figure 5.6: Colony lift filter assay 210
Figure 5.7: Expression and localisation of NPM1 in CRC cells..... 221
Figure 5.8: Immunoprecipitation with NPM1 and EGR1 223
Figure 5.9: Expression and localisation of eEF1A1 in CRC cells 225
Figure 5.10: Immunoprecipitation with EGR1 and eEF1A1 227

Tables

Table 1.1: Hereditary forms of CRC 5
Table 2.1: Cell lines cultured and used within this thesis 64
Table 2.2: Cell line used for DNA and RNA extraction 65
Table 2.3: Mutation and genetic instability status of CRC cell lines 65
Table 2.4: Oligonucleotides for PCR amplification 71
Table 2.5: EGR1 siRNA primers 85
Table 2.6: Antibodies for Western Blot analysis 89
Table 2.7: Antibodies for immunocytochemistry analysis 90
Table 3.1: Genotype of *EGR1* variants in CRC patient samples 106
Table 3.2: *KRAS* and *BRAF* mutations in CRC patients 111
Table 3.3: Wilcoxon rank sum analysis of healthy control tissue after treatment with inflammatory mediators 125
Table 3.4: Wilcoxon rank sum analysis of Ulcerative colitis patient tissue after treatment with inflammatory mediators 130
Table 3.5: Wilcoxon rank sum analysis of Crohn's disease patient tissue after treatment with inflammatory mediators 134
Table 3.6: Genotype of IBD and healthy control patients 137
Table 4.1: siRNA primers 154
Table 4.2: Primary and secondary antibodies 158
Table 4.3: Differences between the IP lysis buffers tested 159
Table 5.1: Summary of the yeast SD media/plates with nutritional additives 199
Table 5.2: Dilutions of Y2H mating plated 200
Table 5.3: Transformation efficiencies of pGBKT7-EGR1 constructs 204
Table 5.4: Positive and negative controls for Y2H screen (Invitrogen) 209
Table 5.5: Potential EGR1 interaction proteins as determined by Y2H screen. 211

Abbreviations

AD	Activating Domain
Ade	Adenine
AKT/PKB	Protein Kinase B
AML	Acute Myeloid Leukemia
ANOVA	Analysis of Variance
AP-1	Activator Protein 1
APC	Adenomatous Polyposis Coli
AR	Androgen Receptor
ARF	Alternate Reading Frame
ATF	Activation Transcription Factor
ATG16L1	Autophagy-Related 16L1
BASP1	Brain Abundant, Membrane Attached Signal Protein 1
BD	Binding Domain
BLAST	Basic Local Alignment Search Tool
BMI	Body Mass Index
BMP	Bone Morphogenetic Proteins
BSA	Bovine Serum Albumin
c-AMP	Cyclic Adenosine Monophosphate
C11orf17	Chromsome 11 open reading frame 17
CARD	Caspase Activation Recruitment Domain
CCGG	Colorectal Cancer Genetics Group
CD	Crohn's Disease
Cfu	Colony-forming unit
CHD4	Chromodomain Helicase DNA-binding protein 4
CIMP	CpG Island Methylator Phenotype
CIN	Chromosomal Instability
CK1	Casein Kinase 1
CLEC2D	C-type Lectin Domain Family 2, member D
CRC	Colorectal Cancer
CREB	cAMP Response Element-Binding
CS	Cowden Syndrome
DAPI	4;,6-diamidino-2-phenylindole
DAVID	The Database for Annotation, Visualisation and Integrated Discovery
DCC	Deleted in CRC
DMBA	7,12-Dimethylbenz(a)anthracene
DMSO	Dimethyl Sulfoxide
DR	Death Receptor
DSS	Dextran Sulphate Sodium
EBS	EGR1 Binding Site
eEF1A1	Eukaryotic Translation Elongation Factor 1 Alpha 1

EGFR	Epidermal Growth Factor Receptor
EGR1	Early Growth Response 1
eIF4A1	Eukaryotic Translation Initiation Factor 4A1
ELK1	Ets-like-1
ERK	Extracellular Signal Regulated Kinases
ESE-1	Epithelial Specific ETS-1
FADD	Fas-Associated Protein with Death Domain
FAP	Familial Adenomatous Polyposis
FBS	Foetal Bovine Serum
FITC	Fluorescein Isothiocyante
FKBP9	FK506 Binding Protein 9
GAIT	Gamma Activated Inhibito of Translation
GI unit	Gastrointestinal unit
GWAS	Genome Wide Association Studies
HC	Healthy Control
HCC	Hepatocellular Carcinoma
HDACs	Histone Deacetylases
HGF	Hepatocyte Growth Factor
His	Histidine
HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HP1	Heterochromatin Protein 1
IAP	Inhibitor of Apoptosis
IBD	Inflammatory Bowel Disease
IFN	Interferon
IGF	Insulin-like Growth Factor
IL2	Interleukin-2
IP	Immunoprecipitation
IRF3	Interferon Regulatory Factor 3
JAK2	Janus Kinase 2
JNK	c-Jun N-terminal Kinases
JPS	Juvenile Polyposis Syndrome
KO	Knock-out
LEF	Lymphoid Enhancer Factor
Leu	Leucine
LHβ	Luteinizing Hormone β
LiAc	Lithium Acetate
LOH	Loss of Heterozygosity
LPS	Lipopolysaccharide
LRR	Leucine-rich Repeat
LRRK2	Leucine-rich Repeat Kinase 2
MAPK	Mitogen-activated Protein Kinase
MCR	Mutation Cluster Region
MDP	Muramyl Dipeptide
MDR1	Multidrug Resistance Protein 1

MLH1	mutL Homologue
MMR	Mismatch Repair
MSH2	mutS Homologue 2
MSH6	mutS Homologue 6
MSI	Microsatellite Instability
NAB1/2	NGF1-A-Binding Protein 1/2
NACA	Mascent Polypeptide-associated Complex Alpha subunit
NAG-1	NSAID-activated Gene-1
NCD1	NAB Conserved Domain 1
NF-κB	Nuclear Factor Kappa-light-chain-enhancer of Activated B cells
NFAT5	Nuclear Factor of Activated T-cells 5
NK	Natural Killer
NLR	NOD-like Receptor
NOD	Nucleotide Oligomerisation
NOD1/2	Nucleotdie-binding Oligomerisation Domain 1/2
NPM1	Nucleophosmin
NSAIDs	Non Steroid Anti-Inflammatory Drugs
NuRD complex	Nucleosome Remodelling and Deacetylase complex
OD	Optical Density
PARP	Poly ADP-ribose Polymerase
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDGF-AB	Platelet-derived Growth Factor-AB
PGE2	Prostaglandin
PGN	Peptidoglycan
PHC1	Polyhomeotic Homolog 1
PI3K	Phosphatidylinositol 3-kinase
PIAS	Protein Inhibitor of Activated Signal Transducer and Activator of Transcription
PIK3CA	Phosphatidylinositol 3-kinase, catalytic, alpha polypeptide
PJS	Peutz-Jeghers Syndrome
PMA	Phorbol 12-myristate 13-acetate
PPARγ	Peroxisome Proliferator-activated Receptor Gamma
PRMT3	Protein Arginine N-methyltransferase 3
PSPH	Phosphoserine Phosphatase
PTEN	Phosphatase and tensin homolog
PTMS	Parathymosin
qRT-PCR	Quantitative Real Time Polymerase Chain Reaction
RIPK2	Receptor-interacting Serine/Threonine-protein Kinase
RP	Ribosomal Protein
RTKs	Receptor Tyrosine Kinases
SAM	S-(5'-Adenosyl)-L-homocysteine
SCID	Sever Combined Immunodeficiency

SDS	Sodium Dodecyl Sulfate
SLC7A8	Solute Carrier Family (amine acid transporter, L-type), member 8
SNPs	Single-nucleotide Polymorphism
SP1	Specificity Protein 1
SREs	Serum Response Elements
SRF	Sreum Response Factor
STAT3	Signal Transducer and Activator of Transcription 3
TAF8	TAF8 RNA polymerase II, TATA Box Binding Protein (TBP)-associated Factor
TCF	T-cell Factor
TF	Tissue Factor
TGF	Transforming Growth Factor
TLR	Toll Like Receptor
TNF	Tumour Necrosis Factor
TPA	12-O-tetradecanoylphorbol-13-acetate
TRAIL	Tumour Necrosis Factor-Related Apoptosis-Induction Ligand
Trp	Tryptophan
UAS	Upstream Activationg Sequence
UC	Ulcerative Colitis
VCAM1	Vascular Cell Adhesion Molecule 1
VEGF	Vascular Endothelial Growth Factor
YAP-1	Yes Kinase-Associated Protein-1
YIF1A	Yip Interacting Factor Homolog A
ZMIZ2	Zinc Finger, MIZ-type containing 2

1 Introduction

Colorectal cancer (CRC) is one of the most common cancers in the UK, and worldwide, and therefore presents a considerable health problem with the incidence rates of CRC increasing. Similarly incidence rates for inflammatory bowel disease (IBD) are also on the rise as much of the world is beginning to adopt an increasingly more western lifestyle. Early Growth Response 1 (EGR1) transcription factor has been identified as a novel susceptibility gene for both CRC and IBD. The research in this thesis will investigate the expression and localisation of the EGR1 transcription factor in colorectal disease in order to have a better understanding of the role that EGR1 may play in both of these diseases.

EGR1 is differentially expressed in several different cancers, and this study will examine the expression of EGR1 in matched normal mucosa and tumour samples of CRC patients, and in IBD patients. It also examines the localisation of EGR1 in colorectal cancer cell lines and investigates some novel EGR1-protein interactions, in particular an interaction between EGR1 and the IBD susceptibility gene product NOD2. There are only a few known proteins that interact with EGR1 and this thesis will investigate some novel EGR1 interactions, determined by a yeast-two hybrid screen.

This chapter will initially introduce colorectal disease, looking at both colorectal cancer and inflammatory bowel disease, which includes Crohn's disease and ulcerative colitis. It will proceed to discuss EGR1 and the role that it has to play in both cancer and inflammation. The chapter will conclude with an overview of some preliminary experiments that preceded this work and an experimental outline for this thesis.

1.1 Colorectal Cancer (CRC)

Colorectal cancer (CRC) is cancer of the colon and rectum. It can also be used to describe cancer of the anus and appendix. It generally develops from adenomatous polyps that form in the epithelial lining of the colon, which then go on to form a carcinoma through a series of genetic mutations. Approximately 60% of tumours are formed on the left side of the colon, mainly the sigmoid colon, rectosigmoid junction and the rectum (Kinzler and Vogelstein 1996).

CRC is the third most common cancer in the UK in both men and women, with the second highest mortality rate after lung cancer (UK National Statistics 2011) and the incidence rates of CRC have greatly increased in the last 30-40 years. In 2008 32,600 men and women were diagnosed with CRC in the UK. The risk of developing CRC is 1 in 16 for men, and 1 in 20 for women (Cancer Research UK 2011) with the risk increasing with age. Worldwide, 60% of all CRC cases occur in developed countries, with America, Australasia and Western Europe having the highest incidence rates. The lowest rates of CRC are seen in Africa (excluding South Africa) and South-Central Asia (Globocan 2008). Eastern European countries and East Asia have seen greatly increasing rates of CRC as they adopt an increasingly more western lifestyle suggesting that there is an important role for environmental factors in the development of CRC.

A family history of the disease is an important risk factor for CRC and approximately 30-40% of CRC is thought to have a hereditary component. However only ~5% of all CRC cases occur through known genetic mutations, which are accounted for by Familial Adenomatous Polyposis (FAP), Hereditary Non-Polyposis Colorectal Cancer (HNPCC) and other rare related disorders. It is evident that both genetic and environmental factors have important roles to play in the development of CRC and both are discussed in detail in the following sections.

1.1.1 Environmental risk factors for CRC

There are several lifestyle factors that have been attributed to the development of CRC. Evidence suggests that a diet of red meat and unsaturated fats, high alcohol consumption, smoking and obesity can all contribute to a risk of developing CRC (Huxley, Ansary-Moghaddam et al. 2009). Patients with diabetes and inflammatory bowel disease (IBD) have also been shown to have a higher risk of CRC. Conversely it has been shown that physical activity, a diet of vegetables and taking non steroid anti-inflammatory drugs (NSAIDs) can decrease the risk of CRC (Potter 1999).

1.1.1.1 Diet

Although it has been suggested that dietary fibre may help prevent against CRC, the data is inconsistent. In some studies there does appear to be a reduced risk of CRC with a consumption of fruit and vegetables (Riboli and Norat 2003; van Duijnhoven, Bueno-De-Mesquita et al. 2009) however others have failed to show a convincing association and at most is likely to have a small inverse association with CRC (Key 2011). One study showed no association with cereal fiber and colorectal cancer (Terry, Giovannucci et al. 2001) but the overall data for this is again inconsistent (Du, Li et al. 2010). A study investigating an association between flavonoids and CRC showed that there an inverse risk on consumption of flavonoids and associations with other dietary components include vitamin B, vitamin D, folate and methionine (Theodoratou, McNeill et al. 2007; de Vogel, Dindore et al. 2008; Theodoratou, Farrington et al. 2008; Theodoratou, Farrington et al. 2008).

The consumption of red meat has been shown to have an association with CRC. This is not surprising when considering that the countries which have the lowest rates of CRC are ones in which a vegetarian and fish diet are prominent, and the omega-3 fatty acid, commonly found in oily fish, has been found to decrease the risk of CRC (Theodoratou, McNeill et al. 2007) . A meta-analysis involving 19 different studies and 8,000 cases showed that there is a 20% increase in risk between individuals who consume the highest amounts of red meat or processed meat to those that consume the lowest (Larsson and Wolk 2006).

1.1.1.2 Physical activity and weight

A review combining 19 different cohort studies showed that exercise is protective against colon, but not rectal, cancer in both men and women (Samad, Taylor et al. 2005). Similarly the study conducted by Huxley et al, 2009, looking at 27 cohorts indicated that physical activity reduced the risk of CRC by 20%, and again it is more protective for colon than rectal cancer. (Huxley, Ansary-Moghaddam et al. 2009), Conversely, a study using 31 cohorts showed that individuals that have a BMI >30kg/m^2 have a 20% greater risk of developing CRC than those with a healthy BMI of <25kg/m^2 (Moghaddam, Woodward et al. 2007).

1.1.1.3 Smoking and alcohol

An analysis using 21 cohort studies showed that there is a significant increase in the risk of CRC based on alcohol consumption. Individuals that are considered heavy drinkers are at 60% increased risk compared with light or non-drinkers (Huxley, Ansary-Moghaddam et al. 2009). Cigarette smoking also significantly increases the risk for CRC, with a greater risk evident for rectal than colon cancer, and the incidence increases with duration of smoking (Liang, Chen et al. 2009).

1.1.1.4 Nonsteroidal Anti-inflammatory Drugs (NSAIDs)

It is known that NSAIDS, such as aspirin, can reduce the risk of developing CRC by up to 40%. A study showed that a low dose of aspirin over a prolonged period of time (5-10 years) significantly reduces the risk of CRC (Din, Theodoratou et al. 2010). The toxicity of NSAIDs renders them impractical for use as a chemopreventative agent for CRC so the mechanism by which they exert their effect is of great interest.

1.1.2 Hereditary forms of CRC

There are several inherited forms of colorectal cancer, the two most common being Familial Adenomatous Polyposis (FAP) and Hereditary Non-Polyposis Colorectal Cancer (HNPCC). Investigations into FAP and HNPCC has allowed for significant progress to the understanding of the development of sporadic CRC. The major forms of hereditary colorectal cancer with the genetic mutations they are associated with are summarised in Table 1. FAP and HNPCC will be discussed in detail below.

Syndrome	Gene	Lifetime risk
Familial Adenomatous Polyposis (FAP)	APC	100%
Attenuated FAP (AFAP)	APC	69%
Hereditary Nonpolyposis colorectal cancer (HNPCC)	MMR genes (esp. MSH2, MLH1)	80%
MUTYH-Associated Polyposis (MAP)	MUTYH	80%
Peutz-Jeghers syndrome (PJS)	STK11	39%
Juvenile polyposis syndrome (JPS)	SMAD4/BMPR1A	39%
Cowden syndrome (CS)	PTEN	rare

Table 1.1: Hereditary forms of CRC

(Migliore, Migheli et al. 2011)

1.1.2.1 Familial Adenomatous Polyposis (FAP)

FAP is an autosomal dominant syndrome that is caused by a germline mutation in the adenomatous polyposis coli (APC) gene (Groden, Thliveris et al. 1991; Nishisho, Nakamura et al. 1991). It occurs in 1 in 12,000 patients, and accounts for ~ 0.5-1% of all CRC cases. FAP is characterised by the development of hundreds to thousands of adenomas, with one or more progressing to carcinoma by 35-40 years of age, predominantly on the left side of the colon (70-80%) (Lynch and de la Chapelle 2003). Homozygous APC mutations in mice are embryonic lethal, and a

heterozygous mutation results in the development of intestinal polyps similar to those found in FAP and are used as a model for intestinal tumorigenesis (Heyer, Yang et al. 1999; Aoki and Taketo 2007).

The APC gene is located on chromosome 5q21 and the majority of APC mutations found in FAP result in a stop codon and a truncated protein. APC has a role in cell adhesion and migration and is also known to be involved in the regulation of the cell cycle. APC is modulated via the Wnt signalling pathway, which is involved in the regulation of β-catenin (Galiatsatos and Foulkes 2006). In the absence of stimulation β-catenin is held in multi-protein complex with APC, Axin, GSK-3p and CK1. The β-catenin molecule can become phosphorylated by protein kinase A (PKA), and other kinases, which leads to its ubiqinitation and degradation. Upon stimulation by Wnt, the β-catenin is no longer held in complex and does not become phosphorylated. It translocates to the nucleus where it activates gene expression. Loss of APC causes accumulation of β-catenin in the nucleus resulting in gene activation of transcription factors T-cell factor (TCF) and lymphoid enhancer factor (LEF) and a constitutively active pathway (Galiatsatos and Foulkes 2006).

APC gene codes for a ~300-kDa protein which is localised in the cytoplasm, at the basolateral membrane of CRC epithelial cells (Smith, Johnson et al. 1993). There are several distinct domains in the APC protein, including a basic domain which binds microtubules, Armadilo repeats, and β-catenin, axin, EB1 and HDLG binding domains, illustrated in Figure 1.1. Germline mutations are mostly found at the 5' half of the coding region, particularly in exon 15, around codons 1061 and 1309 (Polakis 1995). Two-thirds of the mutations found are frameshift mutations with the remaining being single base substitutions. It has been shown that the mutated APC protein does not bind microtubules (Smith, Levy et al. 1994). It was also shown that expression of wild-type APC in CRC cells which endogenously express mutated APC leads to an increase in apoptosis in the cells (Morin, Vogelstein et al. 1996).

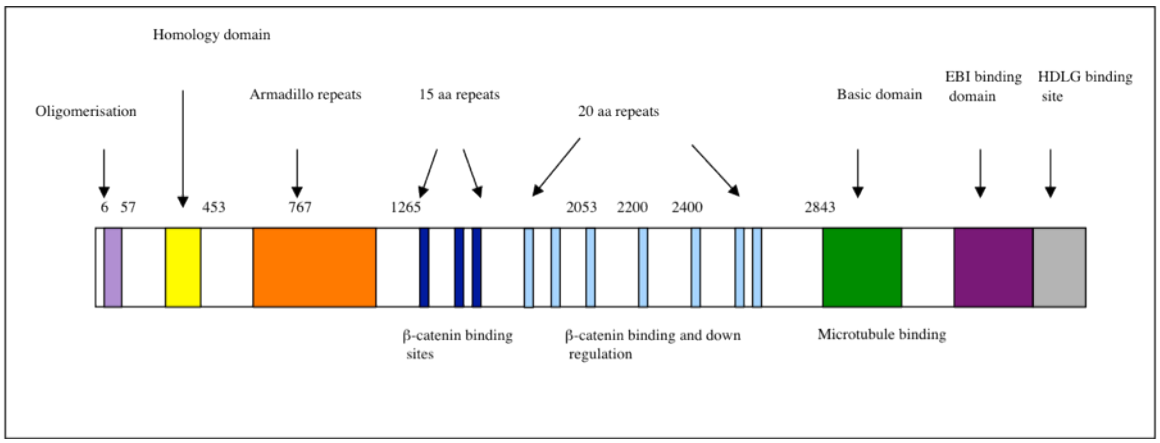


Figure 1.1: Illustration of APC and its protein binding domains.

(Polakis 1995). The APC protein is ~300kDa, with binding sites for β-catenin and mircotubules as well an EB1 binding site and a HDLG binding site.

1.1.2.2 Hereditary Non-polyposis Colorectal Cancer (HNPCC)

Hereditary Non-polyposis colorectal cancer (HNPCC) is the most common form of hereditary CRC accounting for ~4% of all CRC cases. Patients have an earlier onset of CRC than sporadic CRC and the tumours are predominantly localised on the right side of the colon (Lynch and de la Chapelle 2003). HNPCC is caused by germline mutations in DNA mismatch repair (MMR) genes. The MMR systems involves several proteins which are involved in the recognition and repair of nucleotide errors that may occur during DNA replication.

There are 5 MMR genes that have been shown to carry mutations in HNPCC; *MLH1, MSH2, MSH6, PMS1 and PMS2*. Mutations in *MLH1* (mutL homologue) and *MSH2* (mutS homologue) account for >70% of mutations identified in patients and have been detected in ~45-70% of HNPCC patients. Mouse models have been generated for the human honologues of *MutS* and *MutL* and these MMR knockout mice demonstrate DNA repair defects and cancer phenotypes similar to those of patients with HNPCC (Taketo and Edelmann 2009).

HNPCC tumours are characterised by changes in the sequence of repetitive DNA sequences that are localised at microsatellite regions, termed microsatellite instability (MSI), which occurs after the loss of the MMR genes, with ~90% of HNPCC tumour displaying MSI (discussed in more detail in section 1.1.4.2.). However ~60% of patients who fulfil the clinical hereditary criteria for HNPCC do not have a known MMR mutation (Fearon 2011).The clinical hereditary criteria for HNPCC is determined by the Amsterdam criteria. The Amsterdam criteria states that at least three relatives must have a HNPCC-related cancer. These include CRC, endometrial, small bowel, ureter or renal-pelvis. One person should be a first-degree relative of the other two, at least two successive generations should be affected and at least one person should be diagnosed before the age of 50. FAP should also be excluded (Vasen, Watson et al. 1999).

1.1.3 Adenoma-carcinoma sequence of CRC

An adenoma-carcinoma sequence was first postulated in 1988 by Vogelstein et al (Vogelstein, Fearon et al. 1988), and subsequent studies resulted in the development of the adenoma-carcinoma sequence model to explain the progression of CRC from adenoma to carcinoma through alterations in oncogenes and tumour suppressor genes (Kinzler and Vogelstein 1996). Figure 1.2 illustrates the sequence of the development of colorectal cancer incorportating some of the key genetic mutations and the stage at which they occur and the two main forms of genetic instability that are associated with CRC, chromosomal instability and microsatellite instability. Genetic instability associated with CRC will be discussed in the next section. It is known that mutation of the APC gene is an early occurring event in the adenoma-carcinoma sequence, with loss of APC considered a key event in the initiation of development of CRC (Kinzler and Vogelstein 1996).

The adenoma-carcinoma model depicts the sequence of some of the major mutations and genomic effects that are observed in the progression of CRC. Some mutations occur early in the development of CRC, such as APC, and other mutations are only seen in the later stages of CRC tumour development. However not all of the mutations seen in the CRC are responsible for the initiation, progression and/or maintenance of the tumour. Although it has been estimated that there can be ~80 mutated genes in a CRC tumour, it is thought that less then <15 of that mutations seen CRC are considered “drivers” of CRC initiation or progression and then rest of the mutations seen are only “passenger” mutation, ie ones that have occurred as a result of tumorigenesis, but are not responsible for its initiation, progression or maintenance. Some of the mutations are likely to occur somatically, where as others occur as the increase in genetic instability in the tumour increases with tumour development. For example, it has been shown that 18q21.2 region is consistently disrupted in CRC, and it was thought the DCC gene could potentially have a role as a driver in CRC. However it is now more likely that it is the SMAD4 gene that has a key role in tumorigenesis, not DCC as was originally thought (Sjoblom, Jones et al. 2006; Starr, Allaei et al. 2009; Ji, Tang et al. 2010). Other important mutations

highlighted in the adenoma-carcinoma sequence which are now thought to be important driver mutations include KRAS, p53 and PTEN. Several of the key genes associated with CRC will be discussed in detail in section 1.1.5, with further elobration as to how EGR1 relates to some these key genes in cancer discussed in section 1.3.5.

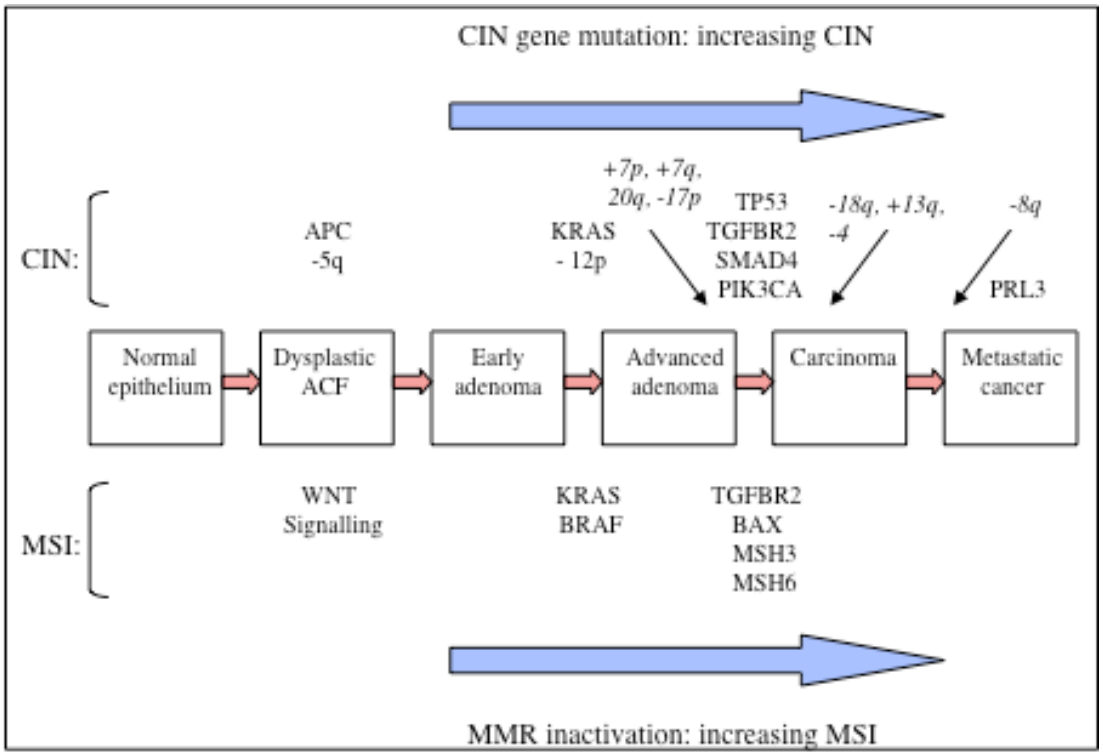


Figure 1.2: Progression of CRC

(Grady and Carethers 2008). Illustration of the adenoma-carcinoma progression of CRC with the associated form of genetic instability, chromosomal instability and microsatellite instability. The key mutations and their timing in the progression of the adenoma-carcinoma sequence are highlighted. Genetic instability and mutations in CRC will be discussed in the following sections.

1.1.4 Genetic instability in CRC

Almost all CRC cases have a form of genetic instability, and it is thought to be an essential step in the formation of cancer. There are two main types of genetic instability associated with CRC; chromosomal instability and microsatellite instability which are mutually exclusive, but result in the loss of the same pathways via different means. Both types of instability occur early in the development of colorectal cancer with the instability increasing as the cancer progresses (Shibata, Peinado et al. 1994; Stoler, Chen et al. 1999) (Figure 1.2).

1.1.4.1 Chromosomal instability (CIN)

Chromosomal instability (CIN) is the most common form of genetic instability in colorectal cancer occuring in ~85% of all CRC cases, and it is this genetic instability that is thought to drive the adenoma-carcinoma sequence (Vogelstein, Fearon et al. 1988; Fearon and Vogelstein 1990). Chromsomal instability is associated with a worse prognosis for CRC (Walther, Houlston et al. 2008). CIN occurs through abnormal segregation of chromsomes and structural re-arrangements such as translocations and deletions which results in the gain or loss of entire chromosomes or large regions of chromosomes leading to aneuploidy and a loss of heterozygosity (Lengauer, Kinzler et al. 1998).

Loss of heterozygosity (LOH) occurs frequently in CRC which can result in the inactivation of tumour suppressor genes and is though to be one of the key steps in the carcinogenesis of CRC. LOH is loss of one allele at a specific point and often occurs at loci where one of the allels is already abnormal e.g. through mutation. For example LOH occurs at chromosome 5q, where it is thought to affect APC, in ~20-50% of CRC tumours and as APC mutations occur at an early stage of CRC development this often results in the loss of both wild-type allele. A large proportion of CRC tumours (70%) show LOH at 17p, affecting p53 which is located at 17p13.1. LOH also occurs at chromosome 18q in CRC tumours (~70%), with evidence to suggest a role for tumour suppressor genes located there such as DCC, SMAD2 and SMAD4 (Grady and Markowitz 2002). These genes are also frequently mutated in

CRC and the relevance of these genes and their signalling pathways will be discussed in section 1.1.5.

The cause of CIN is not yet fully understood as many of the genes causing CIN in tumours are unknown. However studies in yeast have shown that over a 100 different genes can cause CIN in yeast. These include genes involved in cell cycle checkpoint, kinetochore structure and function, centrosome and microtubule formation, chromosome condensation and sister-chromatid cohesion (Grady and Carethers 2008). Interestingly it has also been shown that mutations in APC may contribute to chromosomal instability, through a disruption in the kinetochore-microtubule attachment causing an impairment in chromosome segregation (Fodde, Kuipers et al. 2001; Kaplan, Burds et al. 2001).

1.1.4.2 Microsatellite Instability (MSI)

Microsatellite instability is changes in the sequence of repetitive DNA sequences that are localised at microsatellite regions. Inactivation of MMR genes results in an inability to repair errors that occur during replication at microsatellite sequences. Besides microsatellite regions, DNA replication errors can also occur in genes which have short nucleotide repeats such as A_n or CA_n in their coding region such as *TFGβII, IGF2R* and *BAX*. Tumours can be MSI-high (MSI-H) or MSI-stable (MSS), MSI-stable tumours are usually CIN.

MSI-H tumours appear in ~15% of sporadic CRC cases, which can be caused by mutations in the same MMR genes as HNPCC. About 10% of mutations are in *MSH6* which is rarely mutated in HNPCC (Kinzler and Vogelstein 1996). The criteria of a MSI-H tumour depends on at least two unstable loci out of a possible five loci (BAT25, BAT26, D5S346, D2S123 and D17S250) called the Bethesda panel (Boland, Thibodeau et al. 1998). Inactivation of MMR genes can also occur through aberrant methylation of the gene promoter region and MSI-H tumours in

sporadic CRC are also associated with the CpG island methylator phenotype (CIMP) pathway.

1.1.4.3 CpG island methylator phenotype (CIMP) pathway

The CIMP pathway is characterised by methylation of CpG islands inducing transcriptional silencing of, in general, tumour-suppressor genes, and is thought to be an alternative molecular pathway to CRC than the CIN pathway tumours described earlier (Toyota, Ahuja et al. 1999). Epigenetic changes in DNA occur at CpG dinucleotides, and regions that are enriched for CpG dinucleotides are called CpG islands. These islands usually occur in the 5' regions of genes, and are found at ~50-60% of genes. Global hypomethylation occurs in colorectal neoplasia, at an early stage in the cancer development, resulting in aberrant gene activation. Most CpG islands are unmethylated, however many of these CpG islands can become aberrantly methylated in cancer and this aberant methylation is associated with gene silencing. Aberrant methylation of a gene promoter is an early event in the development of colorectal cancer (Feinberg and Tycko 2004). Methylation of CpG islands occurs more frequently with increasing age and many genes are methylated in normal ageing musosa as well as in tumours. This is referred to as A-type methylation. However a study by Toyota et al, 1999, showed that there are several genes which undergo methylation only in colorectal tumours, known as C-type methylation (Toyota, Ahuja et al. 1999).

In CRC there are many genes which have been found to be aberrantly methylated, including *MLH1*. In CIMP tumours, silencing of the MMR genes is caused by aberrant methylation of predominantly the *MLH1* promoter (Weisenberger, Siegmund et al. 2006) and occurs in >80% of MSI-H tumours. It is also possible to have a CIMP CRC phenotype independent of MSI-H, ie that are microsatellite stable (Jass 2007). The CIMP positive tumours can be split into two subgroups. Group 1 are characterised by MSI, have a high rate of *BRAF* mutations (>50%) and a better prognosis, whereas group 2 are MSS and have a high frequency of *KRAS* mutations

(92%). CIMP positive mutations rarely have mutations in *APC* or *TP53* (Kim, Lee et al. 2010).

The discovery of the CIMP phenotype has led to the idea that there may be three pathways to developing CRC, chromosomal instability (CIN), epigenetic instability (CIMP/MSS) and epigenetic instability with mircosatellite instabiltiy (CIMP/MSI) (Issa 2008). Given the discovery of the CIMP phenotype and a further understanding of the mutations necessary to drive the progression of CRC, Figure 1.3 below, taken from a review by Kim, Lee et al, 2010, has adapted the classical adenoma-carcinoma sequence model and summarised the three potential pathways to sproadic CRC development, with some of the key mutations associated with the different pathways, e.g. the association between BRAF mtuations and the CIMP/MSI pathway.

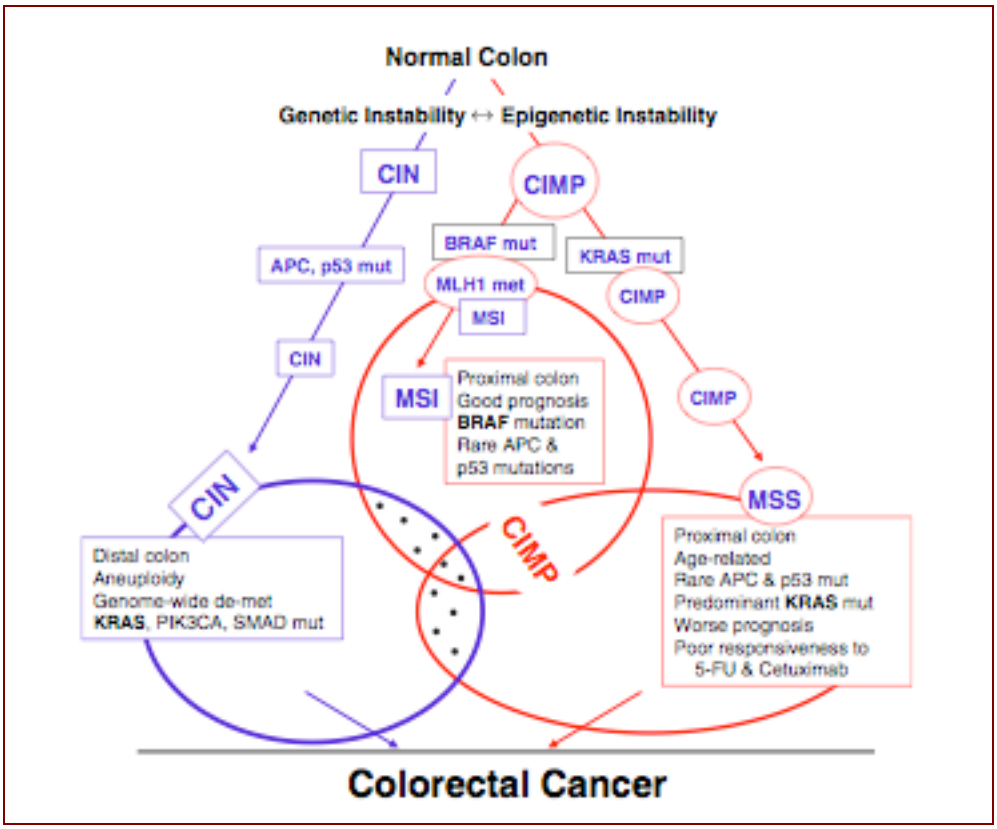


Figure 1.3: Pathways to CRC development

(Kim, Lee et al. 2010). Illustration of the three potential pathways to sporadic CRC development; CIN, CIMP/MSI and CIMP/MSS. CIN tumours are associated with aneuploidy and mutations in *APC*, *p53*, *KRAS*, *PIK3CA* and *SMAD*. CIMP tumours with MSI are associated with BRAF mutations, a good prognosis and methylation of *MLH1* promoter. CIMP tumours with MSS are associated with *KRAS* mutations, have a worse prognosis and cancer development at an increased age.

1.1.5 Mutations in CRC

There are several key genes that are regularly mutated in CRC which lead to its activation and progression including *APC*, *KRAS*, *p53*, *BRAF*, *DCC* and *PIK3CA*. Malignant tumours have a higher number of mutated genes compared with adenomas, suggesting that mutations accumulate through the adenoma-carcinoma development (Lea, Jackson et al. 2009).

1.1.5.1 *APC* mutations in sporadic CRC

Mutations in the *APC* gene resulting in an inactive form of APC, occur in the majority (~70-80%) of sporadic CRC cases (Segditsas and Tomlinson 2006). Mutation of *APC* is an early event in CRC development (Powell, Zilz et al. 1992) (Figure 1.2). Both alleles of APC are inactivated in both sporadic and FAP cancers. Somatic mutations in CRC have been shown to predominantly occur in one region of the *APC* gene termed mutation cluster region (MCR). This cluster region is from codon 1286 to 1513, with mutations at codons 1309 and 1450 most common, similarly to germ-line mutation, 95% of somatic mutations result in a truncated protein product (Miyoshi, Nagase et al. 1992)

Germline mutations in *APC* are 100% penetrant with mutations in *APC* considered to be the initation mutation event in CRC but other mutations are required for development of the early adenocarcinoma to full carcinoma. *APC* mutations occur in the earliest stages of CRC development and it has been postulated that APC is a 'gatekeeper' gene of cell proliferation in the epithelial layer of the colon, and its genetic alteration is enough to initiate tumour formation (Kinzler and Vogelstein 1996).

1.1.5.2 RAS genes

Mutations in *KRAS* are another early event in the development of CRC and are found in 30-50% of tumours (Baldus, Schaefer et al. 2010). The majority of *KRAS*

mutations are located at codon 12 and 13 and result in a constitutively active protein (Davies, Bignell et al. 2002). *KRAS* mutations are thought to contribute to the progression of CRC from early adenoma to an intermediate adenoma but are not necessary for the initiation of CRC development like *APC* (Kinzler and Vogelstein 1996) (Figure 1.2). KRAS is a member of the RAS family, small G-proteins which are regulated by receptor tyrosine kinase growth factors such as EGFR. Upon activation Ras proteins activate downstream signalling pathways such as the MAPK and PI3K pathways (Malumbres and Barbacid 2003). *BRAF* is one of the downstream targets of the Ras proteins and is mutated in ~10% of CRC tumours, predominantly at codon 600 (Davies, Bignell et al. 2002). *BRAF* mutations are associated with CIMP tumours, particularly those that are MSI-H (Domingo, Espin et al. 2004; Kambara, Simms et al. 2004)

1.1.5.3 TP53

The tumour suppressor gene *p53*, located at 17p13.1, is mutated in ~50% of CRC tumours. It is involved in the regulation of DNA damage and normally acts as a tumour suppressor gene. Mutations in *p53* occur in the later stages of the adenoma-carcinoma sequence and are predominantly missense mutations. Mutations in *p53* usually occur at CpG dinucleotide repeats, resulting in G:C to A:T transitions, and effect its DNA-binding activity (Baker, Preisinger et al. 1990).

1.1.5.4 PIK3CA

Mutations are found in *PIK3CA* in ~20-30% of colorectal tumours (Samuels, Wang et al. 2004). They are both constitutents of the PI3K signalling pathway. PIK3CA is an isoform of the p110 catalytic subunit of PI3K. The PI3K signalling pathway is activated via receptor tyrosine kinases (RTKs), which then results in the phosphorylation of AKT and activation of downstream targets. Mutations of PIK3CA usually occur in either the helical or kinase domain (exons 9 and 20), resulting in increased kinase activity and the phosphorylation of AKT in the absence of growth factor stimulation of the RTKs (Chalhoub and Baker 2009).

1.1.5.5 Phosphatase and tensin homolog (PTEN)

Mutations of *PTEN* also occur in CRC (~10%) and are more frequent in MSI-H tumours (Guanti, Resta et al. 2000; Dicuonzo, Angeletti et al. 2001). PTEN is an antagonist of the PI3K pathway and loss of PTEN function results in PIP_3 accumulation, phosphorylation of AKT and constitutive activation of the downstream signalling pathways, which includes activation of NF-κB, and MDM2, the p53 inhibitor (Samuels, Wang et al. 2004; Chalhoub and Baker 2009).

1.1.5.6 TGFβ signalling pathway

LOH occurs at chromosome 18q in CRC tumours (~70%) and is one of the most commonly observed genetic alterations in CRC, with evidence to suggest a role for tumour suppressor genes located there such as *DCC, SMAD2* and *SMAD4*. *SMAD2* and *SMAD4* are mutated in ~5-10% of CRC tumours (Grady and Markowitz 2002). *SMAD2* and *SMAD4* are involved in the TGFβ signalling pathway. TGFβ is a cytokine that is involved in growth inhibition, apoptosis and differentiation. It activates a heterometric receptor complex consisting of serine-threonine kinases. Once the receptor is activated it triggers the phosphorylation of downstream targets including SMAD2 and SMAD3. SMAD2 and SMAD3 can form a heterodimeric complex, or a heterotrimeric complex with SMAD4 and regulate gene transcription involving transcription factors such as tissue factor (TF) (ten Dijke and Hill 2004). Dysregulation of TGFβ signalling can also occur through mutations of *TFGβR2* in ~30% of CRC. In 85% of MSI-H tumours frameshift mutation occur in *TGFβR2*, resulting in the inactivation of the receptor (Markowitz, Wang et al. 1995).

1.1.6 Identification of risk alleles in CRC

Many studies are now focusing on the identification of novel risk alleles for CRC, the hypothesis being that there are many low penetrant risk alleles which convey a small or moderate risk to CRC. A large number of genome wide association studies (GWAS) have been conducted in order to begin to elucidate some of these risk factors. Replicated studies by Dunlop et al have determined up to ten low penetrant risk alleles that confer a risk to colorectal cancer. Interestingly several of these loci contain genes found to be involved in the TGFβ signalling pathway which is already known to dysregulated in CRC. Figure 1.4 below (Tenesa and Dunlop 2009) illustrates a summary of the components of the TFGβ pathways and highlights some of the somatic mutations that are known to occur in CRC as well as the new risk alleles identified. GWAS studies have also identified several susceptibility loci for CRC, with an enrichment for genes linked to the MAPK pathway (Lascorz, Forsti et al. 2010). A meta-analysis of several different GWAS studies identified four novel susceptibility loci for CRC, located at chromosome 1q41, chromosome 3q26, chromosome 12q13 and chromsome 20q13 (Houlston, Cheadle et al. 2010), with several variants identified in the BMP pathway as conferring a risk to CRC (Tomlinson, Carvajal-Carmona et al. 2011). Therefore it may be that there are many other low penetrant risk alleles for CRC in known or novel pathways leading to increased CRC risk.

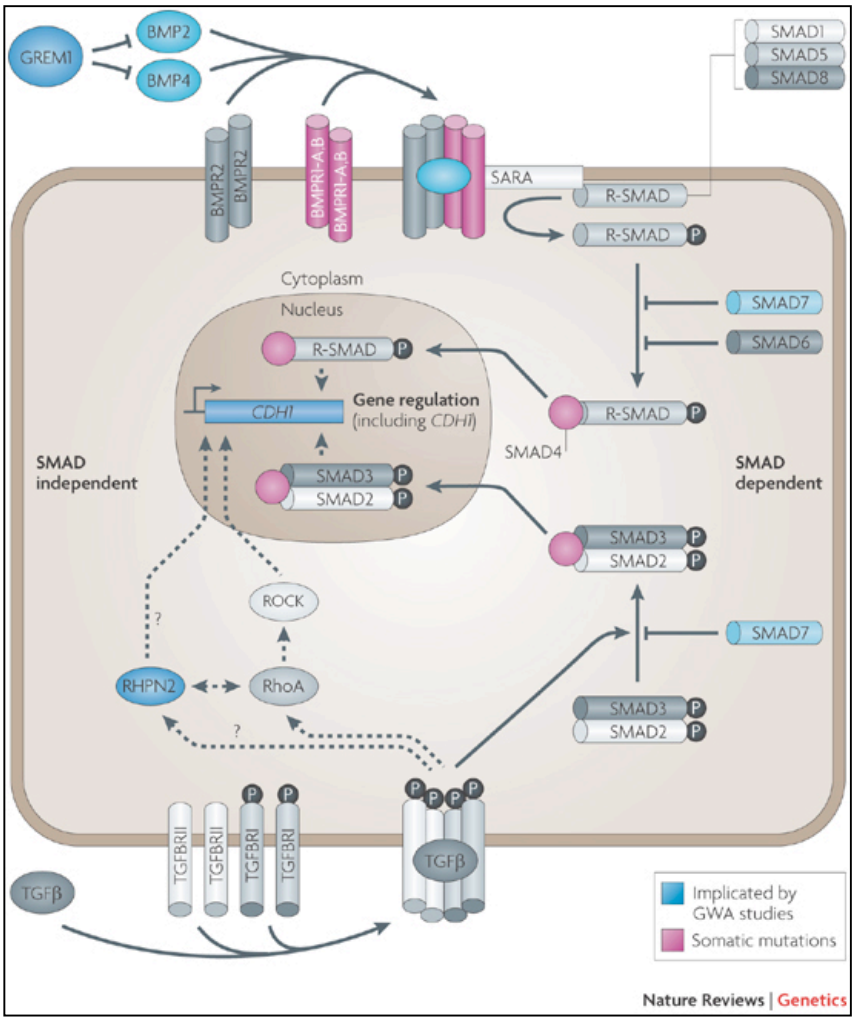


Figure 1.4: TGFβ signalling pathway

(Tenesa and Dunlop 2009). GWAS studies have identified several novel susceptibility loci for CRC, with an enrichment for the TGFβ signalling pathway evident. This figure illustrates the TGFβ pathway with known somatic mutations in CRC in pink, and the potential risk alleles of CRC identified in blue.

1.2 Inflammatory Bowel Disease (IBD)

Inflammatory bowel disease consists of two relapsing disorders, Crohn's disease (CD) and ulcerative colitis (UC), that are characterised by a chronic inflammatory condition of the intestines. IBD has an early onset age, occurring at ~15-30 years of age, with earlier onset occurring in patients with a family history of the disease. Ulcerative colitis presents as continuous inflammation from the rectum extending to the colon, with the inflammation affecting only the mucosal layer of the intestine. However Crohn's disease can affect any part of the gastrointestinal tract, from the mouth to the anus and the disease presents with transmural lesions. The inflamed areas of mucosa are alternated with regions of healthy appearance.

There are high incidence rates of IBD in North America and Northern European countries, with lower incidences reported in Africa, south-east Asia, Australia and South America. This north-south gradient may be due to genetic factors however it is more likely to be due to environmental factors. Incidence rates are increasing in countries that previously had low rates of IBD and there are increasing incidence rates in immigrants from these countries following a move to more developed countries (Vermeire and Rutgeerts 2005; Baumgart, Thomas et al. 2009; Ishihara, Aziz et al. 2009).

IBD is caused by a combination of both environmental and genetic factors, as well as microbial agents and the response of the immune system (Fiocchi 1998). It is known that IBD results from an aberrant response of the immune system to commensal flora and antigens in the intestine. The epithelial layer is the first defense to bacterial infection and acts as a barrier. It has been shown that this barrier can be more permeable in IBD patients with defects found in the junctions of the epithelial cells as well as defects in mucus production (Roda, Sartini et al. 2010). The innate immune response has become of increasing interest in understanding IBD, and several genes involved in the innate immune system have been shown to be associated with IBD (discussed in section 1.2.2).

The immune response of IBD patients is defective in several ways including different expression of Toll-like receptors (TLRs) in intestinal epithelial cells than observed in healthy individuals and defects in T-cell activation and clearance (Vermeire and Rutgeerts 2005; Baumgart and Carding 2007; Ishihara, Aziz et al. 2009).

Although similar, CD and UC have clinical and molecular differences. Mouse models of intestinal inflammation have shown that the inflammation usually involves either a Th-1 or Th2-type T cell mediated inflammation and it is found that CD is characterised by the Th1-type T cell-mediated inflammation, which results in an increase of IL-2, IFN-γ and TNF (Strober, Fuss et al. 2002). In UC however there is an inflammatory response that is associated with a Th2-type T cell inflammation characterised by an increase of IL-4, IL-5 and IL-13 (Bouma and Strober 2003).

1.2.1 Environmental risk factors for IBD

Similar to CRC, IBD is a disease that is attributed to both environmental and genetic risk factors. Environmental factors include smoking, diet, geography, sanitation and hygiene (Fiocchi 1998). Interestingly smoking appears to be only a risk for CD whereas it appears to be protective against UC (Mahid, Minor et al. 2006). Other factors that have been suggested to play a role in the development of IBD include usage of oral contraceptives, appendectomy, stress and childhood conditions.

1.2.1.1 Cigarette Smoking

The evidence for cigarette smoking as a risk factor for IBD is interesting as it has a negative effect on CD patients but a positive effect for UC patients (Mahid, Minor et al. 2006). It has been shown that cessation of cigarette smoking increases the risk UC and patients with UC found that symptoms worsen when they quit smoking, with improvement found when they resumed smoking (Cosnes 2004; Loftus 2004). One study showed that smoking has a dose-dependent beneficial effect on symptoms of UC (van der Heide, Dijkstra et al. 2009). UC patients who smoke are half as likely to require hospitalization for the disease and former smokers with UC are 50% more likely to require hospitalization (Cosnes 2004; Loftus 2004).

Conversely smoking is a risk factor for CD, with smokers more than twice as likely to develop the disease. Former smokers are also at an increased risk compared to people who have never smoked. CD patients who smoke have a higher risk of flare-up, and are associated with more disease complications and a lower quality of life (Cosnes 2004; Loftus 2004).

1.2.1.2 Diet

The effect of diet in the development of IBD has been extensively studied however the results are inconsistent. Several food types have been suggested to play a negative role in IBD, with an excessive consumption of carbohydrates, high consumption of mono-/poly-unsaturated fat, vitamin B and sugar all suggested to

increase risk (Geerling, Dagnelie et al. 2000). Some studies have shown that high consumption of dietary fibre, fruit, vegetables, and fish have been suggested to be protective against the development of IBD, but other studies are inconclusive (Amre, D'Souza et al. 2007).

1.2.1.3 Childhood factors

It has been suggested that children who were breastfed have a lower risk of developing IBD. However although breastfeeding has been postulated to have a protection against IBD, the results are unclear with conflicting studies (Molodecky and Kaplan 2010). It has been shown that CD is more common in patients who had hot water available in the childhood home, but not with UC, with the suggestion that this may be contributing factor to the increase in CD in developed countries in the last 50 years (Gent, Hellier et al. 1994). Other childhood factors that have been suggested to be associated with decreased risk of IBD include growing up in crowded living conditions, large families, absence of tap water, and absence of hot water (Baumgart and Carding 2007; Molodecky and Kaplan, 2010) leading to a 'hygiene hypothesis' where excessive sanitation and hygiene in childhood limits exposure to infections necessary to acquire immunity against diseases. However patients with CD have been shown to be from smaller households with less siblings and lower birth rank has been associated with an increase in the risk of both UC and CD, and potentially due to an increase in the exposure of childhood infections (Hampe, Heymann et al. 2003; Benlloch, Paya et al. 2006; Bernstein, Rawsthorne et al. 2006).

1.2.1.4 Appendectomy

Appendectomy appears to be a risk factor in the development of CD, but confers a protection against the development of UC. It has been shown that there is a negative association between having an appendectomy and the development of UC, with evidence to suggest that an appendectomy is protective for UC (Rutgeerts, D'Haens et al. 1994). However the opposite appears to be evident for the development of CD. A study by Anderson et al, 2003, showed that there is an increased risk of developing

CD, dependent on sex, age and the diagnosis of CD at the time of operation (Andersson, Olaison et al. 2003). An appendectomy is associated with an increased risk of developing strictures in CD (Baumgart & Cardin, 2007), however other studies are less clear as regards to risk with some studies showing no association with risk (reviewed Molodecky & Kaplan, 2010).

1.2.1.5 Other factors

A meta-analysis involving nine studies suggested that there is a modest association between taking oral contraceptives and the development of CD and UC (Godet, May et al. 1995). Although the data is not always consistent it has been suggested that there is an increased risk of IBD in women who take oral contraceptives (Boyko, Theis et al. 1994). Another environmental factor thought to be involved in IBD is stress. It is thought that flare-ups in IBD symptoms can be triggered by stress and depression, with an increased likelihood of relapse in times of adverse life conditions and stress (Mawdsley and Rampton 2005).

A study reviewing the literature and a meta-analysis of 23 studies suggested that a *Helicobacter pylori* infection may be protective against the development of IBD (Luther, Dave et al. 2010), which ties in with the evidence suggesting that children who grow up with more sanitary conditions are more likely to develop IBD (discussed section 1.1.1.3). However the occurrence of IBD is more common after a gastrointestinal infection and patients with IBD have higher levels of mucosal bacteria (Baumgart and Carding, 2007), suggesting a complex relationship between bacterial microbes and IBD.

1.2.2 Genetic risk factors for IBD

1.2.2.1 Family history

A positive family history of the disease is the largest independent risk factor for IBD, there is a strong genetic link for IBD and it has been reported that between 5-22% of all patients with IBD have family member affected with IBD. Hence a family history is a strong risk indicator for IBD and the risk of a first degree relative developing IBD is higher with CD than UC, with siblings being the first degree relative with the highest risk (Russell and Satsangi 2004). It has been shown that there is a relative risk of 13-36% in siblings of CD patients, and a lesser risk in siblings of UC patients of 7-17% (Thompson, Driscoll et al. 1996; Halfvarson, Bodin et al. 2003; Vermeire and Rutgeerts 2005). It is known that monozygotic twins have a higher rate of disease concordance (37% for CD and 10% for UC) than dizygotic twins (7% for CD and 3% for UC). This suggests that genetic factors may play a stronger role in CD than in UC (Thompson, Driscoll et al. 1996; Halfvarson, Bodin et al. 2003; Vermeire and Rutgeerts 2005).

1.2.2.2 Susceptibility genes for IBD

There has been a lot of success in determining susceptibility genes for IBD, using both linkage studies and genome-wide association studies (GWAS). It has been shown that although CD and UC share some susceptibility genes, many are associated with only either form of IBD. In some cases the underlying gene causing the association has been identified (Noble, Nimmo et al. 2006). To date GWAS studies have identified 71 susceptibility loci for CD and 47 susceptibility loci for UC (Anderson, Boucher et al. 2011; Franke, McGovern et al. 2011).

NOD2 was the first susceptibility gene identified for CD using linkage studies and later confirmed by GWAS. Several other genes involved in innate immunity and inflammatory pathways have also been identified such as *IL-23R*, *TLR4*, *TLR5*, *IL-12*, *STAT3* and *NOD1*, whereas other genes identified have been discovered in

interesting and unexpected pathways such as *ATG16L1* involved in the autophagy pathway (Barrett, Hansoul et al. 2008). There have been several knock-out mouse models developed for IBD genes and although some knock-out strains such as *STAT3* KO or *IL-10* KO knock-out can spontaneously develop intestinal inflammation other knock-out mouse models including *NOD2* KO and *ATL16L1* KO do not, suggesting that other factors, either genetic or environmental, are necessary for IBD to develop (Mizoguchi and Mizoguchi 2010).

1.2.2.3 Il-23 pathway

One interesting gene candidate that has been identified in IBD is *IL-23R,* specifically an amino acid polymorphism Arg38Gln. Association at this locus has also been implicated in several auto-immune disorders (The Wellcome Trust Case Control Consortium 2007; Barrett, Hansoul et al. 2008). *IL-23R* is located on chromosome 1q31 and the IL-23 receptor is a heterodimeric membrane receptor to the IL-23 cytokine. It forms in complex with the IL-12RB subunit, which has also been found to be a susceptibility gene for CD. Interestingly two other components of the IL-23 signalling pathway have also been found to be associated with IBD. Janus kinase 2 (JAK2), a proximal kinase and downstream target, and signal transducer and activator of transcription 2 gene (STAT3) are both associated with CD and UC indicating that the IL-23 signalling pathway is important in the inflammation in IBD (Barrett, Hansoul et al. 2008; Ishihara, Aziz et al. 2009) (Figure 1.5). JAK2 and STAT3 are also activated in other signalling pathways, including IL-10, which has also been found to be associated with IBD (Andersen, Ernst et al. 2010).

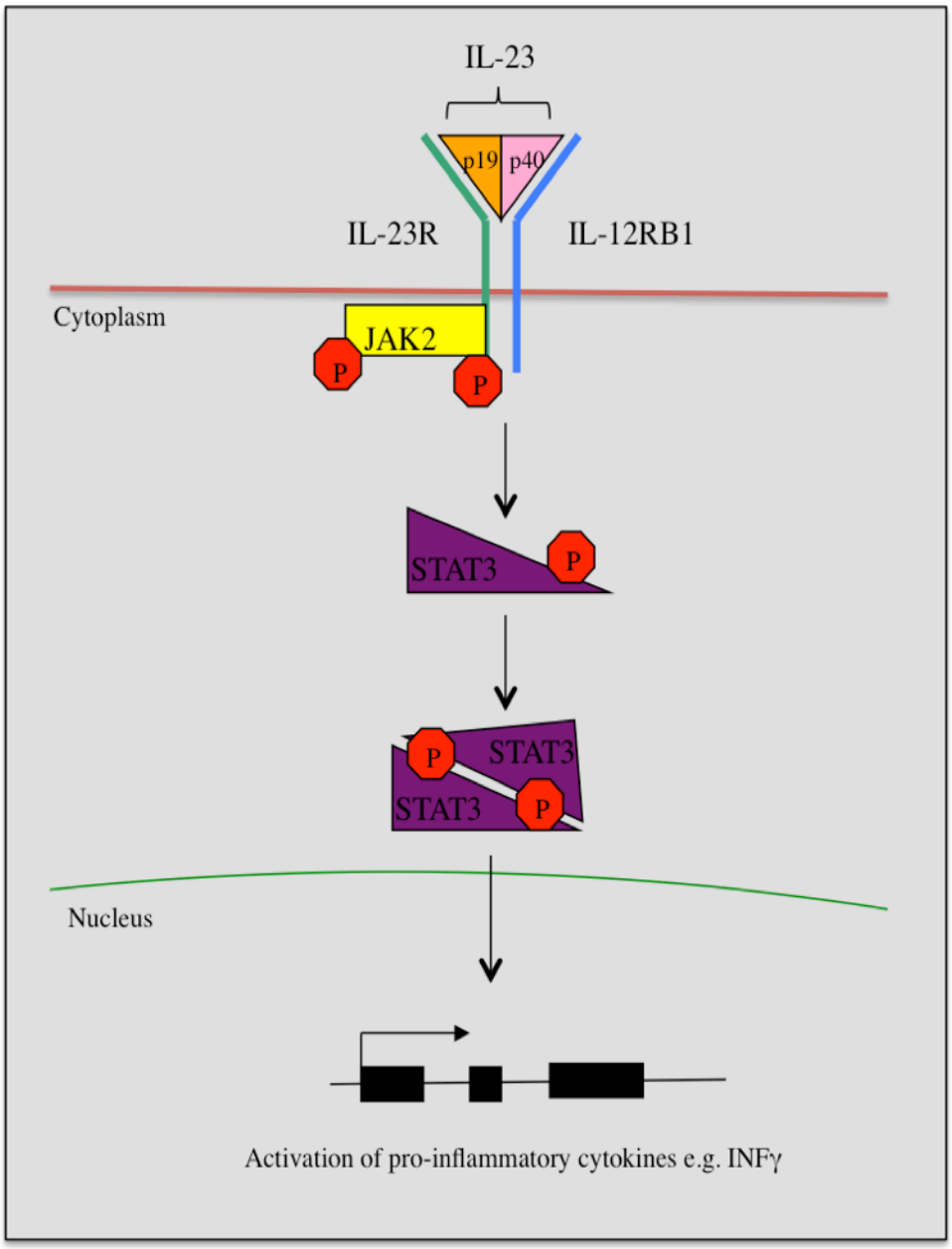


Figure 1.5: Illustration of the IL-23 pathway

(Budarf, Labbe et al. 2009). This figure highlights the genes that are associated with IBD in the IL-23 inflammatory pathway, demonstrating how the use of GWAS studies have identified novel pathways involved in IBD. Gene variants in *IL-23R, IL-12RB, JAK3* and *STAT3* have been shown to be significantly associated with either CD or UC.

1.2.2.4 Autophagy pathway

The finding that IBD is associated with *autophagy-related (ATG) 16L1* gene has generated much interest in the field. It is located on chromosome 2q37.1 and was found that *ATG16L1* is associated with an increased risk for CD, with the strongest association at a non-synonymous amino acid change, alanine to threonine (Budarf, Labbe et al. 2009). It is a key component in the modulation of the autophagy pathway. The autophagy pathway involves enclosing cellular content in the autophagosome, a double membrane organelle, which is then delivered to the lysosome for degradation. It is now thought to play an important role in the innate immune system by delivering intracellular bacteria to the lysosome (Hampe, Franke et al. 2007; Barrett, Hansoul et al. 2008).

Interestingly two other components of this pathway are thought to be associated with CD, *IRGM* located on chromosome 5q33.1 and *leucine-rich repeat kinase 2* (*LRRK2*) located on chromosome 12q12 (The Wellcome Trust Case Control Consortium, 2007; Barrett, Hansoul et al, 2008). NOD1 and NOD2 have been shown to be co-localised with ATG16L1, and it has been suggested that this localisation at the cell membrane is an essential step in the initiation of the autophagy pathway after bacterial infection (Travassos, Carneiro et al. 2010).

1.2.2.5 Nucleotide-binding Oligomerization Domain 2 (NOD2)

CARD15 was the first candidate gene to be identified for IBD. It is located at chromosome 16q12 and polymorphisms in NOD2, encoded by *CARD15*, were found to be significantly associated with CD in western populations (Ogura, Bonen et al. 2001). NOD2 polymorphisms have not been found in Japanese or Asian populations. NOD2 is cytoplasmic protein, which is activated upon recognition of the bacterial ligand, muramyl dipeptide (MDP), a component of both Gram-positive and Gram-negative bacteria. It is involved in the regulation of the inflammatory process and upon MDP recognition, NOD2 activates a signalling pathway that results in the transcription of the inflammatory mediator NF-κB and the activation of the MAPK pathway (Strober, Murray et al. 2006). NOD2 has also been found to be activated by

viral ssRNA, activating IRF3 and production of IFN-β (IL-6) suggesting a role in the immune response to viral infection as well as bacterial (Sabbah, Chang et al. 2009).

NOD2 is a member of the large protein NLR (NOD-like receptor) family which share three structural domains, a leucine-rich repeat (LRR) domain which is involved in ligand recognition, a nucleotide oligomerisation (NOD) domain involved in self-oligomerisation and a caspase activation recruitment domains (CARD) or PYRIN domain that is involved in protein-protein interactions. NOD2 contains two CARD domains located at the N-terminal and is known to be involved in protein-protein interactions with several proteins such as RIPK2. It also contains a central NOD domain which is important for self-oligomerisation that is necessary for NF-κB activation and a C-terminal LRR domain which is required for MDP recognition (Fritz, Ferrero et al. 2006) (Figure 1.6).

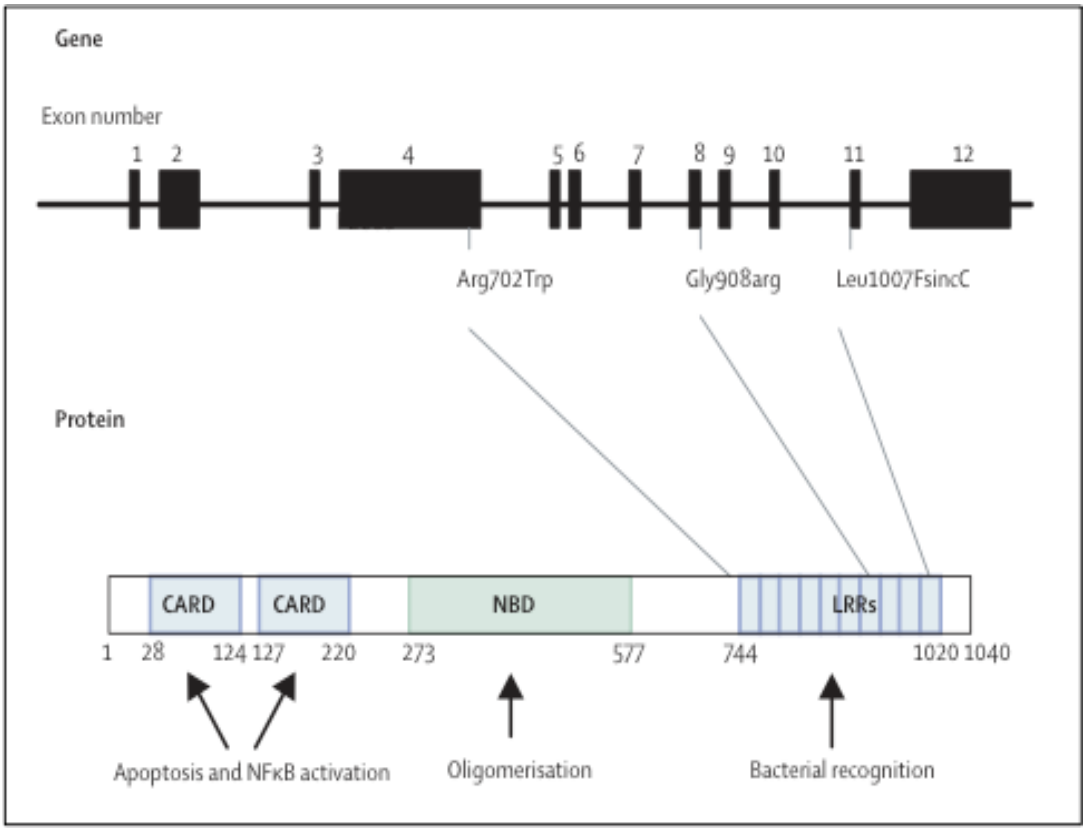


Figure 1.6: Structure of NOD2 gene and protein

(Gaya, Russell et al. 2006). This illustrates the structure of the gene and the protein of NOD2. The gene structure highlights the location of the three NOD2 polymorphisms associated with CD. The protein structure demonstrates the three structural domains of NOD2, two CARD domains, a central NOD binding domain and a LRR domain, along with their associated functions.

MDP activation of NOD2 causes recruitment of receptor-interacting serine/threonine kinase (RICK) via the CARD domain. This results in the activation of the NF-κB pathway and gene transcription. It has been demonstrated in intestinal epithelial cell lines that membrane localization of NOD2 is necessary for the activation of NF-κB in response to MDP recognition (Barnich, Aguirre et al. 2005). Disruption of actin cytoskeleton has been shown to increase NOD2 mediated NF-κB transcription after stimulation by MDP (Legrand-Poels, Kustermans et al. 2007). Activation of NOD2 by MDP can also activate the MAPK signaling pathway through phosphorylation of JNK, p38 and ERK (Abraham and Cho 2006).

The majority of NOD2 mutations are localised in the C-terminal LRR domain, which is involved in ligand recognition (Figure 1.6). There are three mutations associated with CD, Arg702Trp, Gly908Arg and Leu1107fsinsC (3020insC). The Leu1107fsinsC is a frameshift mutation, resulting in a truncated short terminal protein, and tends to carry an increased disease risk than the other two mutations. A study by Maeda et al, 2007, produced a mouse with a NOD2 3020insC mutation. The mice were healthy, with no abnormalities present in the intestines, however they showed that this mutation had a gain of function observed as an increase in NF-κB activation after MDP treatment. After treatment with DSS, the mice displayed an increase in bacteria-induced inflammation and an increase in proinflammatory cytokines, especially Il-1β (Maeda, Hsu et al. 2005). However in human macrophages, the three NOD2 mutations do not appear to respond to MDP stimulation resulting in a decrease in the NF-κB inflammatory response (Tanabe, Chamaillard et al. 2004). It has been shown that the Leu1107fsinsC protein cannot localise to the cell membrane as the deletion causes the loss of the final 33 COOH-terminal amino acid residues, the region that is responsible for the membrane targeting of NOD2 (Barnich, Aguirre et al. 2005). The Leu1107fsinsC protein has also been shown to inhibit the expression of IL-10, an anti-inflammatory cytokine, in human monocyte cells (Noguchi, Homma et al. 2009).

Patients with IBD have an increased risk of developing CRC (discussed in the following section) and one study identified a potential association between the

NOD2 mutation Leu1107fsinsC and an increase in the risk of developing CRC at an older age (Kurzawski, Suchy et al. 2004). A meta-analysis conducted by (Tian, Li et al. 2010) looked at eight studies, consisting of 3,524 CRC cases and 2,364 controls and found that the three NOD2 polymorphisms (Arg702Trp, Gly908Arg and Leu1107fsinsC) are associated with an increased risk of developing CRC.

1.2.3 IBD and CRC

There is strong a link between IBD and CRC, where patients with IBD have a greater risk of developing CRC, which is increased with duration of symptoms and severity of inflammation and dysplasia, and accounts for 15% of all deaths in patients with IBD (Ekbom, Helmick et al. 1990; Munkholm 2003; Jess, Gamborg et al. 2005). Although patients with IBD only account for ~2% of all CRC patients, there is an increased incidence rate of 2.64 for CD and 2.75 for UC patients (Bernstein, Blanchard et al. 2001). A meta-analysis of 116 studies showed that the risk of UC patients developing CRC is ~2% after 10 years of disease, increasing to 8% after 20 years and 18% at 30 years (Eaden, Abrams et al. 2001). Chronic inflammation can promote carcinogenesis, by inducing gene mutation, epigenetic changes, promoting cell growth and angiogenesis. Both positive and negative regulators of inflammation have been implicated in the development of CRC, including TNFα, IL-1, IFN-β, Il-10, COX-2, TLR4, NF-κB and TGFβ (Kraus and Arber 2009; Danese and Mantovani 2010).

The development of colitis-associated CRC has several differences than the models of CRC development discussed in the previous section (Section 1.1). One difference is the stage of cancer in which the *APC* gene is mutated. As discussed previously, the mutation of *APC* is an early event in the development of sporadic CRC and is considered the initiation event of the adenoma-carcinoma pathway. In colitis-associated CRC, it appears that *APC* mutation is a rarer and much later occurring event in the development of the cancer (Kern, Redston et al. 1994; Aust, Terdiman et al. 2002) and mutation of p53, followed by LOH, is an early event (Holzmann, Klump et al. 1998; Hussain, Amstad et al. 2000) (Figure 1.7). Mouse models with *p53* knockout ($p53^{-/-}$), $p53^{+/+}$ and $p53^{+/-}$ do not develop CRC, however after treatment with dextran sulphate sodium (DSS), which causes acute intestinal inflammation, 20% of the $p53^{+/+}$ and $p53^{+/-}$ mice developed cancers which increased to 57% in the *p53* knockout mice and the studies showed that the loss of *p53* enhanced the development of colitis-associated cancer in these models (Chang, Coudry et al. 2007).

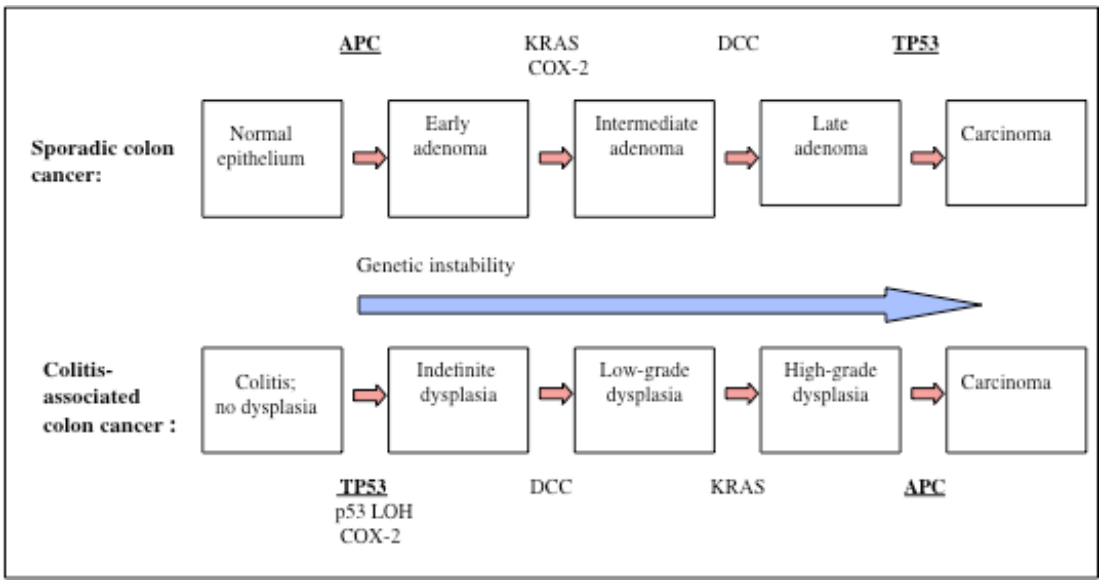


Figure 1.7: Development of sporadic CRC versus colitis-associated CRC

(Xie and Itzkowitz 2008). Illustration of a simple adenoma-carcinoma progression of sporadic CRC compared with a progression model of colitis-associated CRC. As illustrated *APC* mutation is an early event in CRC development, however it occurs at a much later stage in the colitis-associated cancer. Conversely *p53* mutation is a late occurring event in CRC but occurs at the early stages of colitis-associated cancer development.

1.3 Early Growth Response 1 (EGR1)

Early growth response EGR1 (also know as Krox-24, NIGFI-A, TIS8, and Zif268) is a transcription factor first described in 1988 (Sukhatme, Cao et al. 1988). EGR1 is a member of the early growth response (EGR) family, which also contains EGR2, EGR3 and EGR4. They share a conserved DNA binding sequence, a zinc finger motif, with between 81 and 93% homology. *EGR1* is located at chromosome 5q31. Expression of *EGR1* is seen throughout fetal mouse development, however knockout of *EGR1* is not embryonic lethal and mice demonstrate normal growth (Sukhatme 1990; Lee, Tourtellotte et al. 1995). However one study found that EGR1 null mice were smaller in size, with defects in the pituitary glands resulting in a lack of luteinizing hormone (LH) and in the case of the females, defects were also seen in the ovaries. In these EGR1 knockout mice, both sexes were sterile (Topilko, Schneider-Maunoury et al. 1998).

EGR1 was initially recognised as a serum-induced transcription factor involved in cell growth and differentiation. However many different cellular functions have since been attributed to EGR1 as well as its ability to be rapidly and transiently induced by many different environmental signals. EGR1 has become of increasing interest as it was found to be differentially expressed in many different cancers and play a role in the regulation, either directly or indirectly, of several tumour suppressor genes including *p53* and *PTEN* (Thiel and Cibelli 2002). It has also been shown that there is an increase in EGR1 protein levels in the inflamed mucosa of IBD patients (Subbaramaiah, Yoshimatsu et al. 2004).

1.3.1 Structure and promoter of EGR1

EGR1 is a zinc-finger transcription factor whose motif consists of three Cys2-His2 type zinc fingers (Pavletich and Pabo 1991). This region acts as a DNA-binding domain, and binds to a specific GC-rich sequence, 5'-GCGGGGGCG-3' (Christy and Nathans 1989). This sequence is also found in the *EGR1* promoter, suggesting that EGR1 can activate its own gene expression. EGR1 also contains a nuclear localisation signal, which is located in a basic region near the zinc finger motif. This basic region is also conserved in the other members of the EGR family, EGR2 and EGR3 (Gashler, Swaminathan et al. 1993). EGR1 is known to bind to DNA as a monomer.

The EGR1 protein has an inhibitory domain situated between the transcriptional activation domain, which is localised at the amino terminus and the DNA-binding domain (Russo, Sevetson et al. 1995) (Figure 1.8 A). This inhibitory domain binds two transcriptional co-factors, NGF1-A-binding protein (NAB) 1 and 2 (NAB1, NAB2), which block the biological activity of EGR1. NAB1 was first identified as an interacting protein of EGR1 by yeast-two hybrid screen using a mouse embryo cDNA library, where it was found to interact via the inhibitory domain and suppress transcriptional activity of EGR1 (Russo, Sevetson et al. 1995).

The 5' genomic flanking region of the *EGR1* promoter contains five serum response elements (SREs) (Christy, Lau et al. 1988) (Figure 1.8 B). The serum response elements require binding of two transcription factors, the serum respone factor (SRF) and a ternary complex factor Ets-like-1 (ELK1) protein. The *EGR1* promoter also contains a cyclic AMP response element which binds CREB. CREB is constitutively bound to the *EGR1* promoter and the phosphorylation of CREB is also necessary for activation of *EGR1*.The *EGR1* promoter contains a SP1 binding site and a NF-κB binding site as well as an EBS site which can bind EGR1 (Thiel and Cibelli 2002; Tur, Georgieva et al. 2010).

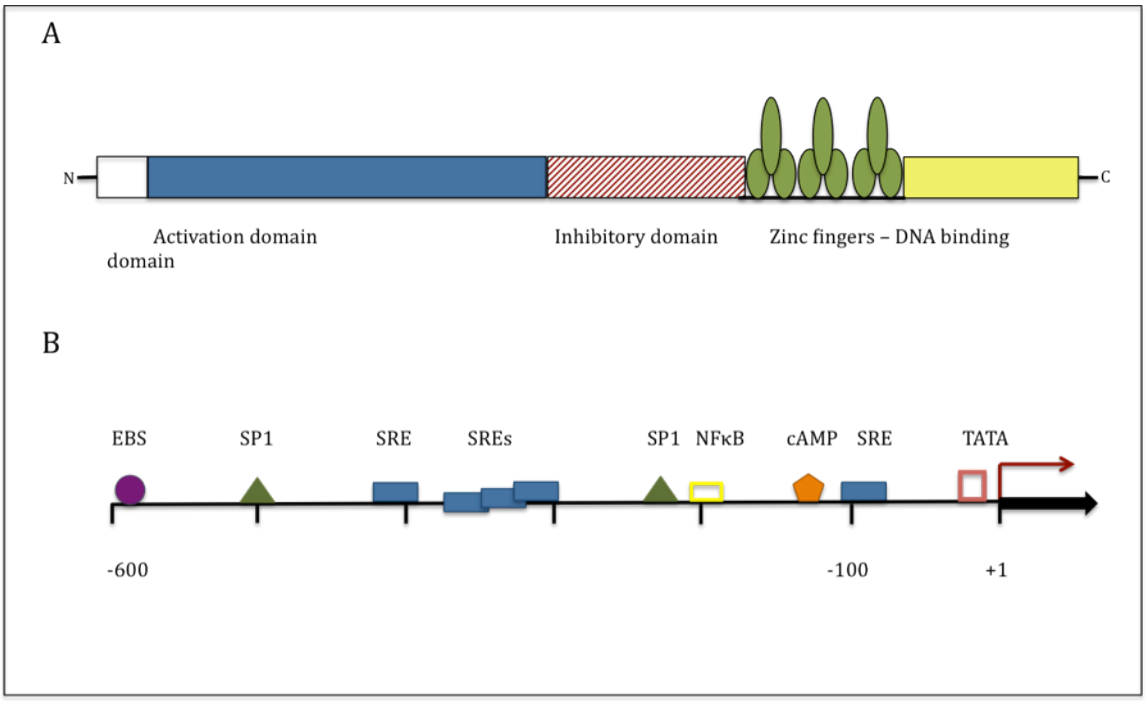


Figure 1.8: Illustration of EGR1 domains and promoter region

(Adapted from Thiel and Cibelli 2002; Tur, Georgieva et al. 2010). **A.** Demonstrates the domains of EGR1, an activation domain, an inhibitory domain which binds two transcriptional co-factors NAB1 and NAB2 and a DNA-binding domain that contains a zinc-finger motif. **B.** An illustration of the *EGR1* promoter with the regulatory binding regions highlighted. The *EGR1* promoter has several serum response elements (SREs), a cyclic AMP site, transcription factor binding regions for SP1 and NF-κB and an EBS region which binds EGR1.

1.3.2 Induction of EGR1 expression

The expression of EGR1 is rapidly induced by many environmental signals, such as growth factors, hormones, cytokines and neurotransmitters, as well as serum and phorbol 12-myristate 13-acetate (PMA) (DeFranco, Damon et al. 1993). It has been shown that activation of EGR1 by serum, PMA, EGFR and LPS is mediated by the mitogen-activated protein kinase (MAPK) signalling pathway, via Ras-Raf-MEK-ERK with ERK1/2 causing phosphorylation of Elk1 which forms a complex with SRF on the EGR1 promoter (DeFranco, Damon et al. 1993; Whitmarsh, Shore et al. 1995; Guha, O'Connell et al. 2001; Maegawa, Arao et al. 2009; Tur, Georgieva et al. 2010) (Figure 1.9).

EGR1 is also induced by stress stimuli such as the chemotherapeutic drugs cisplatin and etoposide, DNA damaging agents, genotoxic stress, UV, and bacteria *Helicobacter pylori* (Virolle, Adamson et al. 2001; Xu, Dziarski et al. 2001; Yu, de Belle et al. 2004; Matsunoshita, Ijiri et al. 2011). It has been shown that induction of EGR1 by growth factors leads to a transient expression, whereas UV-induced expression of EGR1 leads to a sustained expression (Yu, de Belle et al. 2004). The EGR1 promoter has a NF-κB binding site, and EGR1 expression is induced by NF-κB in prostate cancer cells, with full induction of EGR1 by NF-κB requiring the E2F1 transcription factor. Both NF-κB and E2F1 are over-expressed in prostate cancer, as is EGR1 (Zheng, Ren et al. 2009).

The induction of EGR1 by anisomycin, a drug which inhibits DNA and protein synthesis, in T lymphocyte cells is mediated via the MLK2/MKK3/p38 MAPK signalling pathway, which causes the phosphorylated CREB and ATF1 to bind to the EGR1 promoter (Rolli, Kotlyarov et al. 1999) (Figure 1.9). Activation of EGR1 by PGE_2 in macrophage cells is also mediated via p38, with CREB and ATF2 binding to the EGR1 promoter to induce expression (Faour, Alaaeddine et al. 2005). The induction of EGR1 by different stress stimuli can have opposing effects. In prostate cancer it was found that serum-induced EGR1 increases the transcription of p300,

which has several EGR1 binding sites, and this induction of p300 acetylates and stabilises EGR1. However activation of EGR1 by UV irradiation causes the induction of EGR1 phosphorylation, which leads to a downregulation of p300 (Yu, Baron et al. 2007).

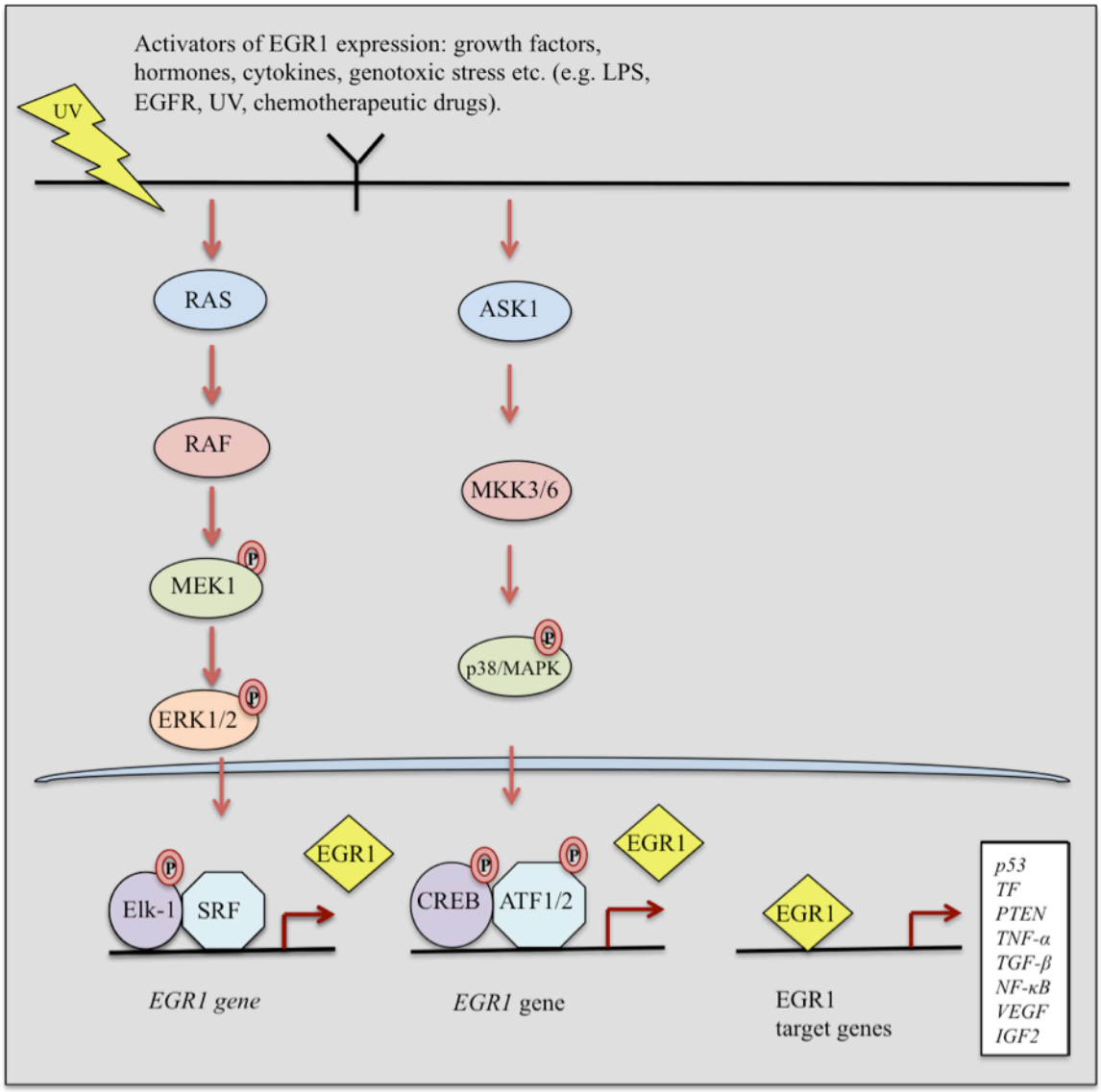


Figure 1.9: MAPK pathways involved in *EGR1* activation

Illustration of the MAPK signalling pathways that regulate the activation of *EGR1*. LPS, PMA and EGFR induction of *EGR1* is known to be mediated via MEK/ERK whereas induction of *EGR1* by PGE_2 and anisomycin involves MKK/p38 signalling.

1.3.3 Regulation of EGR1

As an early response gene the expression of EGR1 is normally rapid and transient, depending on the stimulus and cell type, with mRNA levels of *EGR1* increasing as soon as 15min after activation. The main method of regulating *EGR1* expression appears to be in the form of the two inhibitory co-factors, NAB1 and NAB2, as well as EGR1 itself.

1.3.3.1 NGF1-A-binding protein (NAB)-1 and 2

NAB1, one of the two inhibitory co-factors of EGR1 is ubiquitously expressed and can completely block EGR1-mediated transcription. However *NAB2* expression is induced by the same stimuli as *EGR1*, such as serum and growth factors (Svaren, Sevetson et al. 1996). *NAB2* contains some EGR1 binding sites in its promoter and transcription of *NAB2* is activated by EGR1. It has been shown in epithelial cells that this activation by EGR1 is repressed by NAB2 itself in a dose-dependent manner, i.e. EGR1 activation of *NAB2* causes *EGR1* transcription to be switched off suggesting EGR1 and NAB2 act in a negative feedback loop (Kumbrink, Gerlinger et al. 2005). The expression of NAB2 is also induced by other members of the EGR family, EGR2 and EGR3 indicating that they play a role in the regulation of NAB2 expression (Kumbrink, Kirsch et al. 2010). NAB2 has been shown to block the EGR1-mediated transcription of tissue factor (TF) in HeLa cells as well as preventing the EGR1-mediated production of growth factors TGF1, VEGF, HGF and PDGF-AB (Houston, Campbell et al. 2001).

NAB1 and NAB2 have the ability to multimerise with each other, forming either homo or hetero multimers, through a conserved domain, NAB conserved domain 1 (NCD1). NAB1 and NAB2 contain two highly conserved domains, NCD1, which is also required for interaction with EGR1, and NCD2 which is though to be required for the repression activity of EGR1 (Svaren, Sevetson et al. 1996; Svaren, Sevetson et al. 1998). It has been found that the c-terminal domain of NAB2 interacts with chromodomain helicase DNA-binding protein 4 (CHD4), which is a component of

the nucleosome remodelling and deacetylase (NuRD) complex. This interaction is necessary for the inhibitory function of NAB2 (Srinivasan, Mager et al. 2006).

EGR1 is known to be overexpressed in prostate cancer. It has been shown that *NAB2* expression is down-regulated in 80% of prostate cancers, suggesting that loss of NAB2 results in an unbalanced expression of *EGR1*, with several of EGR1 target genes similarly overexpressed in prostate cancer. The loss of NAB2 in prostate cancer appears to be an early event in the development of prostate cancer (Abdulkadir, Carbone et al. 2001). Interestingly studies have indicated that NAB2 may also play a role as a co-activator of EGR1. It has been shown that NAB2 is a co-activator of EGR1-mediated transcription of luteinizing hormone β (LHβ) in CV-1 cells and of IL-2 transcription in Jurkat cells (Sevetson, Svaren et al. 2000; Collins, Wolfraim et al. 2006).

1.3.3.2 Methylation

It is known that changes in cytosine DNA methlyation are one cause for epigenetic changes in cancer. These changes occur within the dinucleotide CpG. Areas that are rich in CpG regions, called CpG islands, are found in the 5' region of the gene, and are usually unmethyated (Esteller 2007). CpG island hypermethlyation of genes can cause silencing of the gene and it has been found that several tumour suppressor genes are hypermethylated in human cancer, indeed several genes have been found to be CpG methylated in their promoter region in prostate in breast and CRC. Genes that usually are methylated can also become un-methylated in cancer, thereby causing aberrent gene activation (Rodriguez, Frigola et al. 2006; Schuebel, Chen et al. 2007; Chung, Kwabi-Addo et al. 2008; Kakar, Deng et al. 2008; Toyota, Suzuki et al. 2008).

A number of promoters regulated by EGR1 have been found to undergo altered methylation such as the p73 promoter, and the MDR1 promoter in prostate cancer (Enokida, Shiina et al. 2004; Pipaon, Real et al. 2005). It has been found that the heparanase promoter is CpG island hypomethylated in both bladder and prostate cancer (Ogishima, Shiina et al. 2005; Ogishima, Shiina et al. 2005) and the

expression of heparanase is known to be regulated by EGR1 in tumour cells (de Mestre, Rao et al. 2005). The methyaltion of these promoters may affect the binding of EGR1 to its binding site of the promoter which may have an effect on gene activation. *EGR1* has several CpG islands in its 5' region and it has been suggested that CpG methylation of the EGR1 promoter may influence the transcription and expression of *EGR1* (Seyfert, McMahon et al. 1990).

1.3.4 The EGR1 protein and transcription factor

EGR1 codes for an 80kDa nuclear protein that is rich in amino acids proline and serine (~33%) and is endogenously expressed. It is known to have several phosphorylation sites, although these are not well characterised, and the protein is short lived when not phosphorylated. The EGR1 protein is thought to be regulated by the ubiquitin-proteasome pathway as EGR1 can be ubiquinated and has been shown to interact with PRC8, a subunit of the 20S proteasome complex (Bae, Jeong et al. 2002). EGR1 is known to localise to the nucleus and as mentioned earlier it contains a nuclear localisation signal. However there is evidence to suggest that EGR1 may also be localised in the cytoplasm and that the cytoskeleton plays a role in shuttling EGR1 between the nucleus and the cytoplasm (Mora, Olivier et al. 2004).

1.3.4.1 EGR1 and the Cytoskeleton

It has been suggested that the cytoskeleton may play a role in the regulation of EGR1 in benign prostate cells as rapid nuclear translocation of EGR1 is mediated by the cytoskeleton in these cells (Mora, Olivier et al. 2004). Curcumin, the active component of the herb *Curcuma longa,* was shown to inhibit cellular proliferation and trigger cell death partly by binding to tubulin, causing tubulin aggregation and perturbed microtubule assembly in epithelial cells (Gupta, Bharne et al. 2006). Curcumin is known to inhibit EGR1 by suppressing the induced expression of *EGR1* mRNA resulting in decreased expression of EGFR in colon cancer cell lines (Pendurthi and Rao 2000). Conversely, studies using rat hippocampus and primary neurons culture has shown that overexpression of EGR1 promotes the phosphorylation of Tau resulting in the destabilisation of microtubules in the brain (Lu, Li et al. 2011).

1.3.4.2 EGR1 protein interactions

There are not many proteins that are known to interact with EGR1. The most well characterised EGR1 protein-protein interaction is with NAB1 and NAB2 as discussed. The function of EGR1 as a transcription factor is well characterised and the known EGR1-protein interactions appear to facilitate EGR1 in this role. EGR1 has been shown to interact with Yes kinase-associated protein –1 (YAP-1), which is a transcriptional coactivator of p73, in prostate cancer cells. After irradiation, this EGR1-YAP1interaction induces Bax gene expression causing apoptosis in the cell (Zagurovskaya, Shareef et al. 2009). As mentioned earlier serum induction of EGR1 in prostate cancer cells causes the protein to become acetylated, after which it is able to directly bind to p300/CBP where it induces expression of growth genes such as TGFβ, and IGF-2, and inhibits transcription of itself, and p300/CPB. Conversely induction of EGR1 by UV causes the protein to become phosphorylated, after which it transactivates pro-apoptotic genes *p53*, *PTEN* and *Bcl-2* (Yu, de Belle et al. 2004; Adamson, Yu et al. 2005).

1.3.4.3 EGR1 target genes

EGR1 has been shown to activate many different genes which are involved in many different and even opposing cellular functions including genes involved in growth and differentiation, cell cycle regulation, apoptosis and tumour promoting as well as trannscription factors. Some interesting target genes of EGR1 include the growth factors and cytokines *IGF2, TGFβ,* and *VEGF* and apoptotic proteins such as p53, PARP, PTEN and TNF-α. EGR1 activates expression of the cell cycle regulatory protein Cyclin D1, extracellular matrix proteins fibronection and VCAM1 as well other transcriptional regulatory proteins such as Fos, ATF3, PPARγ, NF-κB and EGR1 itself (Adamson and Mercola 2002; Fu, Zhu et al. 2003; Krones-Herzig, Mittal et al. 2005). EGR1 regulates genes that have been shown to be important in human cancer as well as genes that are involved in the immune and inflammatory response. The role of EGR1 in cancer and the inflammatory response will be discussed in detail below.

1.3.5 EGR1 and cancer

EGR1 expression has been found to be significantly lower in a variety of cancer cells and human tumours including breast carcinoma, glioblastoma, osteogenic sarcoma, basel cell carcinoma, squamous cell carcinoma, fibrosarcoma, uterine leiomyomas and esophageal carcinoma (Huang, Liu et al. 1995; Huang, Fan et al. 1997; Calogero, Arcella et al. 2001; Wu, Chen et al. 2001; Pambuccian, Oprea et al. 2002; Fang, Wee et al. 2007). In contrast, *EGR1* seems to be overexpressed in gastric cancer and the levels of *EGR1* expression correlated with the clinical stage of the cancer as well as the tumour infiltration and invasion of the tumour (Zheng, Pu et al. 2010). The levels of *EGR1* are also significantly higher in prostate cancer, and the increased levels of *EGR1* may play a role in influencing the regulation of genes including signalling proteins, transcription regulators, neuroendrocine proteins and membrane-associated proteins involved in adhesion and signalling (Svaren, Ehrig et al. 2000). It has been suggested that EGR1 may act as either a tumour promoter or as a tumour suppressor as both gain and loss of *EGR1* appears to be significant in human cancer.

Many tumour cell lines also express little EGR1 and re-introduction of EGR1 into these cell lines, such as HT1080 (a fibrosarcoma cell line), inhibits the growth and transformation of the cells and induces expression of TGFβ1 (Liu, Yao et al. 1999). Similarly re-introduction of EGR1 to esophageal carcinoma cell line and tissue resulted in a reduced growth, reduced soft agar colony formation and reduced tumour growth rate in SCID mice (Wu, Chen et al. 2001). In an *egr1*$^{-/-}$ mouse model, introduction of a Lewis lung carcinoma (LLC1) showed slower rates of tumour growth compared with wild-type mice (Caso, Barry et al. 2009). In a study by Krones-Herzig et al, in which a two-stage carcinogenesis experiment using DMBA and TPA was performed, it was shown that the induced tumours in EGR1 null mice developed significantly earlier than *egr1*$^{+/+}$ or *egr1*$^{+/-}$ mice. This demonstrated an increased susceptibility for tumour development in the absence of EGR1 (Krones-Herzig, Mittal et al. 2005). These studies indicate that there is a difference in the role of EGR1 in tumour development between introducing the mutation into an EGR1

null mouse and inducing a tumour in an EGR1 null mouse, suggesting that the difference in the function of EGR1 depends on the cellular or tumour environment.

One mechanism by which EGR1 is thought to exert tumour suppressor or tumour promoter effects is by its regulation of several tumour suppressor genes, which are known to be involved in the induction of apoptosis. Apoptosis is an important mechanism in regulating cell growth, and the dysregulation of the apoptotic pathway is a regular occurrence in cancer. EGR1 has been proposed to have a role in inducing apoptosis in cancer via three main pathways; binding to the c-Jun transcription factor; transactivation of the *PTEN* gene; and through binding to the p53 promoter with activation of the p53 tumour suppressor gene (Thiel and Cibelli 2002). Both p53 and PTEN are frequently mutation in cancer, and CRC, as previously discussed. The c-Jun transcription factor has been shown to be essential for apoptosis in neuronal cells (Ham, Eilers et al. 2000).

1.3.5.1 EGR1 and p53

TP53 is an essential tumour suppressor gene due to its role in regulation of cell growth and regulation of many mediators of the apoptotic pathway. Activation of p53 results in cell cycle arrest, apoptosis and senescence. Loss of p53 results in dysregulation of the apoptotic pathway and aberrant cell growth. EGR1 has been shown to regulate the expression of p53 and p73 which in turn can regulate the expression of EGR1, which has p53 responsive elements in its promoter region indicating a system of feedback loops that comes into effect in response to stress, resulting in prolonged expression of the p53 family of genes and apoptosis in tumour cells (Yu, Baron et al. 2007). There are several potential EGR1 binding sites in the promoter region of p53, and EGR1 binds directly to the p53 promoter (Krones-Herzig, Mittal et al. 2005).

It has been shown that UV-induced expression of EGR1 only occurred in $p53^{-/-}$ cells. EGR1 expression was not induced in cells with p53 expression, either in $p53^{+/-}$ or $p53^{+/+}$ cells. Over-expression of EGR1 in $p53^{+/+}$ caused transformation of the cells (Zhang and Chen 2001). However in $egr1^{-/-}$ mice, irradiation treatment does not

induce expression of either p53 or p73, indicating that EGR1 is necessary for the induction of p53 and p73 in response to radiation (Yu, Baron et al, 2007).

It has been suggested that the mutation status of p53 may have an effect on the function of EGR1 again indicating the importance of the cellular environment on the function of EGR1. EGR1 can induce the expression of wild-type p53 and may enhance its anti-apoptotic and growth inhibitory effects. However in cells expressing mutant p53, EGR1 appears to have a tumour-promoting role and mutant p53 can induce the expression of EGR1. Loss of EGR1 in cells expressing mutant p53 increases the rate of apoptosis after genotoxic stress, suggesting that expression of EGR1 in cells expressing mutant p53 contributes to its resistence to apoptosis (Weisz, Zalcenstein et al. 2004). Studies by Sauer et al suggest that the over-expression of EGR1 that occurs in prostate cancer may be caused by the presence of mutant p53. In prostate cells, mutant p53 activates the MEK/ERK signalling pathway, which regulates the transcription of EGR1. Activated EGR1 then activates the EGFR/ERK signalling cascade, causing a positive feed back loop (Sauer, Gitenay et al. 2010). Figure 1.10 below illustrates the proposed model of the feedback loop.

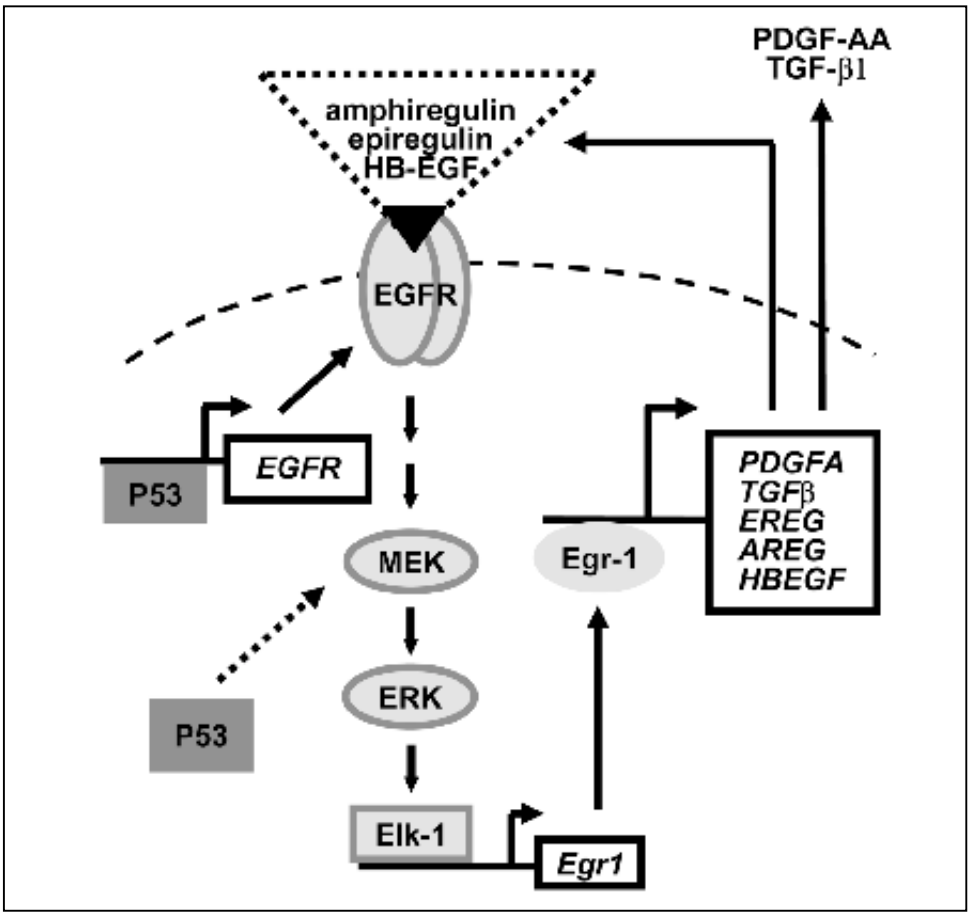


Figure 1.10: Positive feeback loop with EGR1 and mutant p53

(Sauer, Gitenay et al. 2010). Mutant p53 inatiates activation of the loop by induction of MEK, which activates trancription of EGR1, and allows activation of EGR1 target genes. EGR1 up-regulates the expression of EGFR ligands, which activates the EGF receptor resulting in activation of p53.

1.3.5.2 EGR1 and PTEN

PTEN is a lipid phosphatase that acts as a tumour suppressor gene by antagonising the PI3/Akt pathway (Cao, Wei et al. 2004). The *PTEN* promoter has several putative EGR1 binding sites, and it has been shown that EGR1 can directly transactivate the PTEN gene. A study by Virolle et al demonstrated that the presence of EGR1 is necessary for the induction of PTEN following radiation and etoposide treatment, as *egr1*$^{-/-}$ cells are unable to induce expression of PTEN, as well as being resistant to UV induced apoptosis. Induction of EGR1 by UV light results in the expression of PTEN as well as apoptosis. As EGR1 expression is lost in many cancers it has been suggested that EGR1 may play a part in resistance to radiation treatment (Virolle, Adamson et al. 2001).

It has also been shown that induction of ARF by IGF-1 can cause phosphorylation of EGR1 and its ultimate translocation to the nucleolus. The resulting ARF-mediated sumoylation of EGR1 increases the transcriptional ability of EGR1 to induce transcription of PTEN. Like PTEN the ARF promoter contains EGR1 binding sites, with EGR1 able to induce expression of ARF. PTEN expression is not induced in *egr1*$^{-/-}$ or *Arf*$^{-/-}$ cells (Yu, Zhang et al. 2009).

1.3.5.3 EGR1 and colorectal cancer

A RNA microarray study showed that seven genes, including *EGR1*, were consistently up-regulated in the normal mucosa of CRC patients compared with the mucosa from healthy controls (Hong, Ho et al. 2007), indicating that *EGR1* expression is up-regulated in colorectal cancer patients. EGR1 was also shown to be differentially expressed in different stages of colorectal cancer development between normal and adenoma tissue, and between adenoma and carcinoma also (Habermann, Paulsen et al. 2007).

The role of apoptosis in the development of cancer is well known, with the ability of cancer cells to inhibit or evade apoptosis a hallmark of cancer progression. The pro-

apoptotic role of EGR1 depends on the cell type and the nature of the stimulus used to induce EGR1 expression. Several studies have been undertaken to try and elucidate the apoptotic activity of EGR1 in epithelial and CRC cells. Survivin is an inhibitor of apoptosis (IAP) protein that is abberantly expressed in human cancer and over-expressed in several cancer cell lines including the CRC cell line SW480. It has been shown that EGR1 regulates the expression of survivin and binding of EGR1 to the survivin promoter downregulates its expression (Wagner, Schmelz et al. 2008).

As well as mediating the induction of the proapoptotic proteins p53 and PTEN as discussed, EGR1 also mediates the induction of the proapoptotic proteins ATF3 and NSAID-activated gene-1 (NAG-1). Treatment of CRC cells lines (HCT116, SW480, LoVo and HT-29) with the NSAID, tolfenamic acid, suppressed the proliferation of these cells and induced apoptosis through an increase in the nuclear accumulation of epithelial-specific ETS-1 (ESE-1) transcription factor, resulting in the activation of EGR1, which in turn transactivated the expression of the pro-apoptotic protein NAG-1 (Lee, Cho et al. 2008).

Activation transcription factor 3 (ATF3) is a downstream target of EGR1, which has also been shown to be act as a tumour suppressor in colorectal tumours. ATF3 appears to mediate LY29002-induced apoptosis in human colorectal cancer cell lines, HCT-116, SW480, Caco-2 and HT-29 and contrary to PTEN events this apoptosis pathway is a PI3K/Akt-independent pathway (Yamaguchi, Lee et al. 2006). Tumour necrosis factor-related apoptosis-induction ligand (TRAIL) induces the expression of EGR1 in gastric and CRC cell lines. TRAIL induces apoptosis in cells via binding to the death receptor 4 (DR4) and DR5. However DR4 and DR5 can also mediate induction of pro-survival and inflammatory signalling and EGR1 expression is induced after the induction of DR4 and D5R. The DR5-induced apoptosis pathway is inhibited by c-FLIP, which is a target gene of EGR1, indicating an anti-apoptotic role for EGR1 (Mahalingam, Natoni et al. 2010).

Hypoxia inhibition of EGR1 prevents upregulation of EGFR, EGR1 directly induces EGFR transcription and it is also known that low oxygen levels and EGFR play a

crucial role in tumour development and progression (Nishi, Nishi et al. 2002). Treatment of CRC cell lines HT-29 and Caco2 with curcumin was also shown to result in downregulation of EGFR by inhibiting ERK signalling which significantly decreased EGR1 gene expression at both a transcriptional and translational level (Chen, Xu et al. 2006).

1.3.6 EGR1 and inflammation

EGR1 gene expression is regulated by both growth stimulatory and inhibitory cytokines including TNF-α, TNF-β, interleukin-1, IFN-α, and IFN-β (Cao, Guy et al. 1992; Granet and Miossec 2004). EGR1 is induced after treatment with the bacterial components lipopolysaccharide (LPS) and peptidogylcan (PGN), which is mediated by MEK/ERK signalling pathway, through binding to Elk-1 and SRF to the SRE-1 domain in the EGR1 promoter (Xu, Dziarski et al. 2001). The induction of EGR1 by LPS in macrophages is rapid and transient, with mRNA levels increased after 30min of treatment, and protein levels expressed after 1-2 hours (Coleman, Bartiss et al. 1992). TNF-α/β, IFN and IL-1 induced expression of EGR1 is similarly rapid and transient (Cao, Guy et al. 1992; Granet and Miossec 2004). EGR1 is also required for LPS-mediated induction of both TF and TNF-α, both of which are also mediated via MEK and ERK (Shi, Kishore et al. 2002). EGR1 is thought to bind to the TF promoter with AP-1 and c-Rel/p65 transcription factors in order to induce gene expression of TF (Figure 1.11). EGR1 is also involved in the PGN-induced expression of TNF-α, which requires the EGR1, c-Jun and NF-κB transcription factors (Xu, Dziarski et al. 2001).

EGR1 is important in the LPS-mediated induction of IL-6, which is reduced in $egr1^{-/-}$ cells. It has been demonstrated that although EGR1 is not necessary for the initial expression of many inflammatory mediators, it is required for the sustained expression of these mediators (Pawlinski, Pedersen et al. 2003). Although the majority of EGR1 research in relation to the inflammatory response is conducted in monocytic and macrophage cells, epithelial cells do express the receptor for LPS, toll-like receptor 4 (TLR4) and LPS has been demonstrated to induce EGR1 expression in epithelial cells. EGR1 has been shown to co-localise with TF in kidney and lung epithelial cells, and *egr1*$^{-/-}$ cells show a reduced expression of TF (Pawlinski, Pedersen et al. 2003).

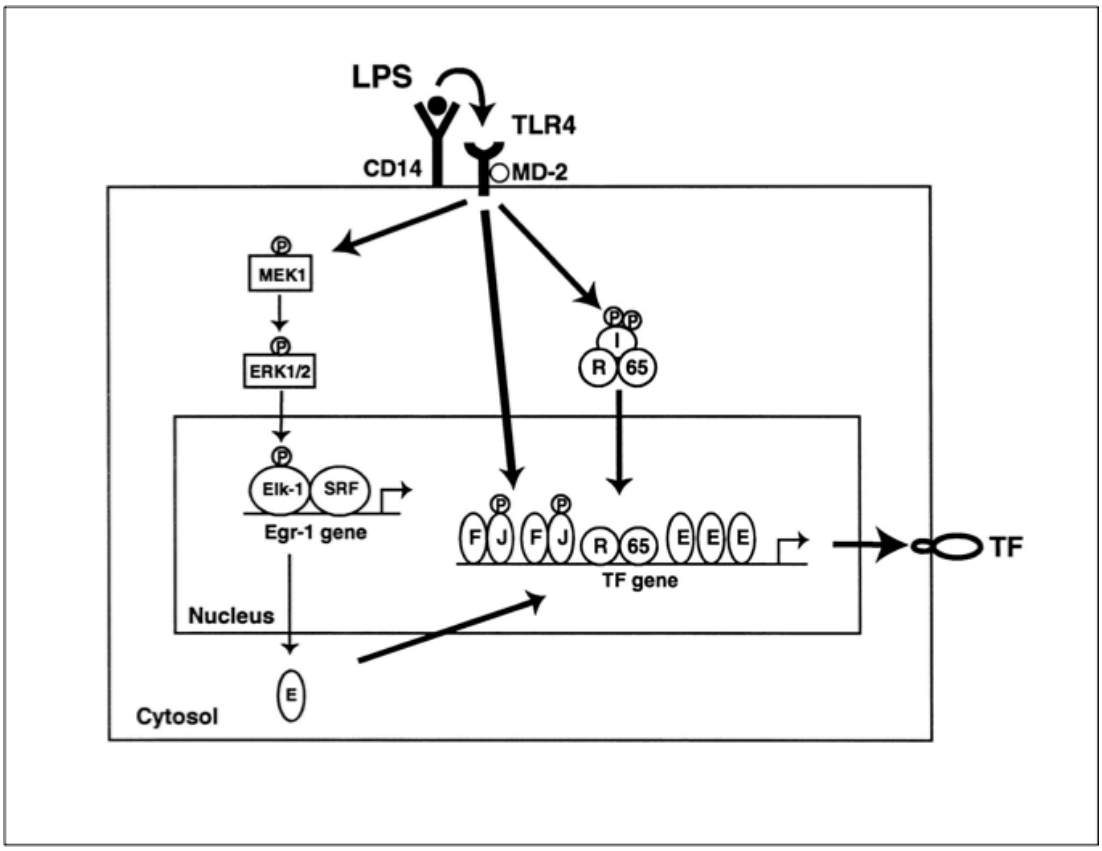


Figure 1.11: LPS induction of EGR1 and TF in monocytes

(Guha, O'Connell et al. 2001). Illustration of the mechanism by which LPS induces the expression of TF via the induction of EGR1. LPS induces the expression of EGR1 via the MEK/ERK signalling pathway which then activates the expression of TF along with transcription factors c-Rel and p65 (labelled R and 65 respectively).

Membrane-associated prostaglandin (PG) E_2 synthase (PGE_2) is an inducible enzyme that is a mediator of inflammation, and is induced by pro-inflammatory stimuli. EGR1 binds to the mPGES promoter, resulting in induction of PGE_2 (Naraba, Yokoyama et al. 2002). Dimethylcelecoxib (a derivative of the NSAID celecoxib used for treatment of rheumatoid arthritis and FAP) inhibits mPGES-1 transcription by down-regulating EGR1 (Deckmann, Rorsch et al. 2010). PGE_2 can induce expression of EGR1 via the MLK2/MKK3/p38 MAPK signalling pathway in a concentration-dependent manner, in macrophage and fibroblast cells. The induction of EGR1 was mediated via binding of ATF2 and CREB transcription factors to the EGR1 promoter. PGE_2 suppressed the induction of TNF-α expression, with EGR1 binding to the TNF-α promoter suppressing transcription (Faour, Alaaeddine et al. 2005).

Helicobacter pylori has been shown to induce EGR1 expression in colon and gastric cell lines, and EGR1 regulates the expression of CD44, ICAM-1, and CD95L (APO/Fas), which are induced after infection with *H.pylori* (Abdel-Latif, Windle et al. 2004), suggesting a role for EGR1 in the innate immune system. EGR1 also contributes to inducible expression of the multidrug resistence-1 (*MDR1)* gene, which codes for P-glycoprotein, and in hematopoietic cells it has been suggested that EGR1 is necessary for MDR1 promoter activity after induction by 12-O-tetradecanoylphorbol-13-acetate (TPA) (McCoy, Smith et al. 1995). MDR1 deficient mice develop large bowel inflammation and dysplasia after infection with *Helicobacter* (Annese, Valvano et al. 2006). A further link with inflammation is the observation that in IBD patients, there was an increase in the levels of EGR1 protein in the inflamed mucosa compared with the non-inflamed mucosa (Subbaramaiah, Yoshimatsu et al. 2004).

1.4 Preliminary studies of EGR1

EGR1 is highly likely to play a role in carcinogenetic processes involved in the development of colorectal cancer. As the above reviewed literature illustrates, EGR1 plays an important role in cancer, as a regulator of either tumour suppressor or tumour promoter genes and also in the induction of apoptosis in cancer cells by various pathways. EGR1 appears to play a significant role in inflammatory pathways, which highlights a role in IBD also. Hence several preliminary studies were undertaken in the Colon Cancer Genetics Group (CCGG) (University of Edinburgh) and the Gastrointestinal unit (GI) (University of Edinburgh) before undertaking this thesis.

In silico studies conducted by the CCGG identified that *EGR1* expression is decreased in the tumour mucosa of patients with CRC compared with normal colonic mucosa (James Prendergast/Susan Farrington pers comm). An in silico analysis of tissue expression datasets was conducted, using 329 adult human EST libraries. The EST libraries contained a minimum of 250 sequences, were grouped according to disease and tissue type, and examined for differences in gene expression between normal and disease, e.g. normal colon and tumour colon. The study showed that EGR1 expression was significantly higher in normal colon (q value = 8.77e-7) and also in normal lung (q value = 4.36e-7). EGR1 expression was not significantly differentially expressed in kidney, liver or prostate tumours.

A co-expression and gene ontology enrichment study was also performed. This study used the GNF Gene Atlas V2 Human U133A data (Su, Wiltshire et al. 2004). The 158 Gene Atlas arrays consisting of 79 tissues (with 2 replicates of each) were normalised and the expression values were calculated using the PM/MM difference model in dChip. The expression values of replicate arrays were pooled and Pearson correlations were calculated between logged values. A funcational annotation term enrichment of the 407 probes displaying the highest co-expression with EGR1 were analysed using DAVID. The following five probes showing high co-expression with EGR1 were mapped to genes that are known to be involved in IBD; Tumour necrosis

factor 2 (TNF2), interleukin 1 beta (IL-1β), solute carrier family 11, intereferon gamma receptor 1 (INFGR) and intercellular adhesion molecule 1 (CD54).

Simultaneously, experiments conducted by the GI unit examined the expression of *EGR1* in colonic biopsies and also showed that *EGR1* is differentially expressed between non-inflamed and inflamed mucosa. A yeast-two hybrid screen conducted by the GI unit using NOD2 as bait against a SW480 CRC cell line library also identified EGR1 as a potential novel interacting protein with NOD2 in SW480 CRC cells.

In collaboration a large scale case-control study was conducted by both the CCGG and the GI unit which demonstrated an association between common variants of *EGR1* with disease phenotype in both CRC and IBD. Three common variants of EGR1 (rs3813321, rs11743810, rs11748288) were genotyped in healthy controls (n=3011) and colorectal cancer patients (n=3012). Overall the study showed that 2 of the EGR1 variants gave a borderline significance for association with CRC (rs3813321 p=0.2967; rs11743810 p=0.0571; rs11748288 p=0.0815) and all 3 EGR1 variants were associated with early onset (<55 years of age) colorectal cancer (rs3813321 p=0.0392, rs11743810 p=0.0283, rs11748288 p=0.0258). Subsequently the 3 variants were genotyped in IBD patients (n=990) and different healthy controls (n=368). This study showed a significant association with 2 of the EGR1 loci for IBD (rs3813321 p=0.0789, rs11743810 p=0.0329, rs11748288 p=0.0413), with a significant association at all three EGR1 loci for ulcerative colitis (rs3813321 p=0.0394, rs11743810 p=0.0202, rs11748288 p=0.0062). Only one of the loci was significantly associated with Crohn's disease (rs3813321 p=0.1882, rs11743810 p=0.0357, rs11748288 p=0.1855) (Susan Farrington/Elaine Nimmo pers comm).

The preliminary data collected by the CCGG and GI unit and reviewed literature suggests that *EGR1* is a candidate susceptibility gene in colorectal disease, both in CRC and IBD, and thus further investigation into the activity of EGR1 in CRC and IBD is warranted.

What is clear from the literature is that EGR1 itself is neither a tumour suppressor nor a tumour promoter gene. What makes it an important gene to study in cancer is its role as a regulatory of gene expression of both tumour promoter and tumour suppressor genes, and of inflammatory genes in both cancer and IBD. As discussed, it has the ability to activate the transcription of many types of genes often with opposing functions, but what determines which genes it activates in still unclear. Therefore it may be that the cellular environment has an impact on the role of EGR1 in these cells. However the mechanism of regulation of EGR1 expression in these cells is unknown. It is therefore necessary to determine first of all to what extent there is differential expression of *EGR1* in colorectal cancer, and to try and determine by what means the expression of *EGR1* is regulated. What is also evident is that the EGR1 protein also can have different roles to play, depending on which stimulus it is induced by ie it is context-dependent. This warrants a further look into EGR1 expression and localization in CRC cells and any potential novel protein interactions in these cells, especially given the potential interaction between EGR1 and NOD2.

1.5 Aims and Experimental Design

The aim of this project is to better understand how EGR1 is expressed in colorectal disease and to further investigate the localisation and function of the EGR1 protein in CRC cells line. This project was conducted in collaboration with the CCGG and the GI unit.

There is little really known about the role of EGR1 in CRC. Therefore we will first determine the expression levels of *EGR1* in matched normal mucosa and tumour samples and this thesis will attempt to determine if EGR1 expression is gained or lost in CRC. Similarly although EGR1 is known to be involved in the regulation of many different inflammatory mediators little is known about its role in IBD. We will investigate the expression of *EGR1* in the normal mucosa of IBD patients along with its expression after stimulation with inflammatory mediators to determine if there is differential expression of *EGR1* in IBD. This will lead to a greater understanding of *EGR1* expression and potential role in both colorectal cancer and IBD.

The mechanisms by which *EGR1* is differentially expressed in cancer is still unclear We will attempt to determine if the *EGR1* variants that were identified as being associated with disease phenotype in both CRC and IBD show any correlation with the expression of *EGR1* in both diseases. Given that EGR1 expression is mediated via the MAPK signalling pathway, and mutation in KRAS/BRAF are a frequent occurrence in CRC, we will determine if mutations in this pathway demonstrate any correlation with EGR1 expression. We will also investigate whether *EGR1* expression is regulated via methylation of the EGR1 promoter in both the CRC and IBD patients.

The finding by the GI unit that EGR1 may interact with NOD2 by yeast-two hybrid experiments is an exciting and novel interaction for both proteins and we will determine if an interaction does occur in CRC cells as this interaction may be important in understanding the role that EGR1 plays in inflammation and IBD. The cellular localisation of EGR1 will be investigated in CRC cells to determine if EGR1

does localise to the cytoplasm, which would allow it to co-localised and interact with NOD2 in the cell. The finding of an EGR1-NOD2 interaction led to the possibility that EGR1 may have many un-identified protein-protein interactions, so a further yeast-two hybrid screen will be conducted to determine novel interacting proteins of EGR1 specifically in CRC cells. It is hoped that this experiment will determine if EGR1 does interact with different proteins and lead to a better understanding of its function in cancer cells.

All experimental techniques used in this thesis are described in detail in Chapter 2. Chapter 3 will discuss the findings in relation to the differential expression of EGR1 in CRC and IBD patients using qRT-PCR as well as sequencing of the three EGR1 variants and bisulfite sequencing of the EGR1 promoter in these patients. Chapter 4 will investigate the expression and localisation of EGR1 in CRC cells using protein extract, Western blotting techniques and immunocytochemisstry and determine if EGR1 interacts with NOD2 and components of the cytoskeleton using immunoprecipitation. Chapter 5 will discuss the yeast-two hybrid screen to determine novel interacting proteins of EGR1.The yeast two hybrid screen is conducted using a SW480 cell line library, and a full length and truncated EGR1 protein.

2 Materials and Methods

This chapter details the methods that are used in this thesis. Where required, more detailed methods are included in the relevant results chapters. Solutions and media denoted with an asterisk (*) have been prepared by the technical services department at the MRC Human Genetics Unit. Where stated, patient DNA and RNA samples have been prepared by members of the CCGG and GI unit.

2.1 Biological Material

2.1.1 Materials and Solutions

Tissue Culture Medium

Dulbecco's modified Eagle's DMEM (Gibco BRL, Invitrogen, UK)

Leibovitz L-15 (Gibco BRL)

McCoy's 5A (Gibco BRL)

10% w/v Foetal bovine serum (FBS)*

1% w/v Penicillin and Streptomycin (P/S)*

Phosphate Buffered Saline (PBS)*

0.1M $NaH_2PO_4.H_2O$

0.1M $NaHPO_4.7H_2O$

pH 7.4

Trypsin Versene (T/V)*

50% w/v Trypsin

50% w/v Versene

Freezing Media

10% w/v Dimethyl Sulfoxide (DMSO) in FBS

Treatments

1mg/ml stock Lipopolysaccharide (LPS; Sigma-Aldrich)

10mM stock Curcumin (Sigma-Aldrich)

2.1.2 Cell culture

Cell lines are stored in the liquid nitrogen facility at MRC Human Genetics Unit. Cells were raised from liquid nitrogen by thawing at 37^{o}C and resuspended in appropriate media (Table 1). Cells were grown in the appropriate media, which was supplemented with 10% foetal bovine serum (FBS)* and 1% penicillin/streptomycin*, and incubated in 5% CO_2 at 37^{o}C. The cells were maintained in the growing phases by removing the media, washing with PBS and treating with a solution of trypsin/versene (1:1) at 37^{o}C for 5min to detach the cells. The cells were resuspended at a suitable dilution in an appropriate amount of media and placed in a fresh flask containing media. To preserve the cell line stock, the cells were centrifuged after they have been detached and the pellet re-suspended in freezing media. The cells were first frozen at –70^{o}C, before transfer to the liquid nitrogen storage at –140^{o}C.

The cells were harvested for RNA, DNA and protein extracted by removing the media and washing the cells in cold PBS. The cells were detached from the flask using a cell scraper and collected in 1ml of PBS. The cells were centrifuged using an Eppendorf Microcentrifuge 5415 R, the PBS was discarded and the cell pellet was used for RNA, DNA or protein extraction immediately or stored at –70^{o}C until needed.

Cell line	Cell type	Media	Reference
HRT18	Colorectal adenocarcinoma	DMEM	(Tompkins, Watrach et al. 1974)
HT29	Colorectal adenocarcinoma	McCoy's 5A	(Fogh, Fogh et al. 1977)
HCT116	Colorectal carcinoma	McCoy's 5A	(Brattain, Fine et al. 1981)
SW480	Colorectal adenocarcinoma	L-15	(Leibovitz, Stinson et al. 1976)
VACO425	Colorectal carcinoma	DMEM	(McBain, Weese et al. 1984)

Table 2.1: Cell lines cultured and used within this thesis

The following cell lines are maintained within the CCGG and cell pellets for RNA and DNA extraction were provided by M. Walker.

Cell line	**Cell type**	**Reference**
SW48	Colorectal adenocarcinoma	(Leibovitz, Stinson et al. 1976)
LoVo	Colorectal adenocarcinoma	(Drewinko, Romsdahl et al. 1976)
CACO2	Colorectal adenocarcinoma	(Fogh, Fogh et al. 1977)
MCF7	Breast adenocarcinoma	(Soule, Vazguez et al. 1973)
PNT	Prostate adenocarcinoma	(Cussenot, Berthon et al. 1991)
PC3	Normal prostate	(Kaighn, Narayan et al. 1979)
DUI45	Prostate carcinoma	(Stone, Mickey et al. 1978)
MDC6	Lymphoblastoid	Lab stock
ConA	Lymphoblastoid	Lab stock

Table 2.2: Cell line used for DNA and RNA extraction

The genetic instability status and mutations where known of the colorectal cancer cell lines used in this thesis are listed below (Din, Dunlop et al. 2004); S.Farrington pers comm).

Cell line	APC	B-catenin	P53	MMR	CIN
HRT18	Mutant	Wild type	Mutant	Deficient (hMSH6)	Negative
HT29	Mutant	Wild type	Mutant	Proficient	Positive
HCT116	Wild type	Mutant	Wild type	Deficient (hMLH1)	Negative
SW48	Unknown	Unknown	Wild type	Deficient (hMLH1)	Unknown
SW480	Mutant	Wild type	Mutant	Proficient	Negative
LoVo	Mutant	Unknowm	Wild type	Deficient (hMSH2)	Positive

Table 2.3: Mutation and genetic instability status of CRC cell lines

2.1.3 Treatment of cells

Cells were grown in T25cm^2 flasks until ~70% confluent and treated with LPS in a dose and time dependent manner, or with 1ug/ml of LPS, MDP, PGN and TNF (Sigma; provided by Elaine Nimmo, GI unit) for 24hours.

2.1.4 Patient Material

2.1.4.1 Colorectal normal mucosa and tumour

Normal and tumour mucosa samples were collected by Prof. Malcom Dunlop. All RNA and DNA extractions on the CRC patient samples were conducted by Dr. Rebecca Barnetson, as detailed in the relevant sections below.

2.1.4.2 Inflammatory bowel disease

The IBD biopsy samples were prepared by the GI unit. The biopsy samples were collected in PBS containing 1% P/S and 1% Gentamycin. The samples were placed on gauze in solution of Weymouths media (Gibco BRL) containing 1% penicillin/streptomycin and 1% glutamine and treated with inflammatory stimuli (LPS, MDP, PGN and TNF). The dish containing the samples was placed over a dish containing H_2O, sealed, and incubated at 37^oC for 24hours. The samples were frozen in RNALater® (Applied Biosystems). All DNA and RNA extractions on the IBD patient material, and cDNA synthesis was conducted by members of GI unit, as detailed in the relevant sections below.

2.2 RNA protocols

2.2.1 RNA extraction

Cell line RNA and CRC patient RNA (extracted by Dr. Barnetson) was extracted using TRIzol reagent (Invitrogen) according to the manufacturer's instructions, and stored at –80^{o}C. IBD patient RNA were extracted by the GI unit using AllPrep DNA/RNA/Protein Mini Kit (Qiagen).

2.2.2 Estimation of RNA concentration

RNA concentration was estimated using optical densitometry on GeneQuant Pro RNA/DNA calculator UV spectrometer (Amersham Pharmacia biotech, Cambridge, UK). RNA samples were diluted 1 in 50 with H_2O, and the absorbency was measured at 260nm and 280nm, providing concentration in μg/ml. RNA samples were considered sufficiently dissolved with a $A_{260/280}$ ratio of >1.8.

2.2.3 DNAse I treatment

Cell line and patient RNA was DNAse treated in a 10μl reaction containing 1μg of RNA, 1μl of DNAse1 (2 Units) and 1μl of 10x reaction buffer at 37^{o}C for 30 min. The reaction was terminated by adding 1μl of DNAse stop and incubated at 65^{o}C for 10 min.

2.2.4 cDNA synthesis

DNAse treated cDNA was converted to cDNA using the 1st strand cDNA Synthesis kit for RT-PCR (AMV), (Roche) as per the manufacturer's instructions. The cDNA samples were stored at –20^{o}C. IBD patient cDNA was prepared by the GI unit.

2.3 DNA protocols

2.3.1 Materials and Solutions

10x Tris-Acetate EDTA (TAE)*

2M Tris

5.7% w/v Glacial acetic acid

50mM Na_2EDTA (pH 8.0)

Loading Buffer

100mM Na_2EDTA (pH 8.0)

0.25% w/v Bromophenol blue

30% w/v Sucrose

2.3.2 DNA extraction

2.3.2.1 Cell Lines

Cell line DNA was extracted using a QIAamp DNA mini kit as per manufacturers instructions by myself and Marion Walker.

2.3.2.2 Patient Material

Colorectal normal and tumour patient DNA was extracted by Dr. Barnetson using using TRIzol reagent (Invitrogen) according to the manufacturer's instructions, and stored at –20^{o}C. IBD patient DNA was extracted by the GI unit using AllPrep DNA/RNA/Protein Mini Kit (Qiagen).

2.3.2.3 Bacterial plasmids

DNA from bacterial plasmids was extracted using a QIAprep miniprep kit or a QIAprep maxiprep kit (QIAGEN), depending on the concentration of DNA required, according to the manufacturers instructions.

2.3.3 Purification of DNA

DNA was purified using QIAquick PCR purification kit (QIAGEN) according to manufacturers instructions.

2.3.4 Estimation of DNA concentration

DNA concentration was estimated using optical densitometry on GeneQuant Pro RNA/DNA calculator UV spectrometer (Amersham Pharmacia biotech, Cambridge, UK). DNA samples were diluted 1 in 50 with H_2O, and the absorbency was measured at 260nm and 280nm, providing concentration in μg/ml. DNA samples were considered sufficient with a $A_{260/280}$ ratio of >1.6

2.3.5 Bisulfite treatment of DNA

DNA samples (1μg) were bisulfite treated using the EpiTect Bisulfite Kit (Qiagen Ltd), as per the manufacturer's instructions.

2.4 PCR protocols

2.4.1 Oligonucleotides

Primers were supplied by Sigma as precipitates and resuspend in dH_2O to a stock concentration of 100μM, with working stocks diluted to 20μM with dH_20. PCR amplification was performed on a Peltier PCT225 thermal cycler (MJ research, Waltham USA) using the following conditions, unless indicated otherwise: 94°C - 3min, (94°C - 45secs, 58°C - 45secs, 72°C - 45secs) x35, 72°C - 10min. All primers used are listed in table 3 below, with F indicating forward and R for reverse.

Primer	**Sequence 5'-3'**	**Cycling conditions**	**Reference**
EGR1-E57F EGR1-E57R	CAAGAGTTAGGGCTGGGACAT GCTTCTTTGATGCCCTCAAG	Anneal at 65°C	Farrington pers.comm.
EGR1-E60F EGR1-E60R	CCAGGAACATTGCAATGTGC GTAACTGGTCTCCACCAGC	Anneal at 55°C	Farrington pers.comm
EGR1-E61F EGR1-E61R	CTGCTCAGTTCGTGCTCACT CCCCTTTATCCAGTCCAGGT	Anneal at 60°C	Farrington pers.comm.
BRAF-F BRAF-R	GATATTGCACGACAGACTGCAC CTGGTCCCTGTTGTTGATGTTTG	As standard	Novel
KRAS-Exon2-F KRAS-Exon2-R	ACTGAATATAAACTTGTGGTAGTT GG CGAGAATATCCAAGAGACAG	As standard	Novel
KRAS-Exon3-F KRAS-Exon3-R	ACGATACAGCTAATTCAGAATC CATAGGTACATCTTCAGAGTC	As standard	Novel
KRAS-Exon4-F KRAS-Exon4-R	CTTTGTGTATTTGCCATAAATAAT AC CACACAGCCAGGAGTCTTTTC	As standard	Novel

EGR1-Reg1-F1 EGR1-Reg1-R EGR1-Reg1-F4 (Nested)	TTTAGGGATTGTAGGGGGAGT ATACTACCCCCTAAAAATTCCCTCT TTTTTTTGGGGTGAGTAATTGATT	Anneal at 60^{o}C 65^{o}C (Nested)	MethPrimer (Li and Dahiya 2002)
EGR1-Reg2-F1 EGR1-Reg2-R1 EGR1-Reg2-F2 (Nested)	TTTGGGTAGTATTTTATTTGGAGTG ACCTCCATCCTACACCCTAAAAC GTTTGGGTTTTTTTAGTTTAGTTTA	Anneal at 60^{o}C 65^{o}C (Nested)	MethPrimer (Li and Dahiya 2002)
EGR1Fl-F EGR1Fl-R	AGATC**GAATTC**TGCTCGTCCAGGA TGGCC GGTCC**GAATTC**TTAGCAAATTTCA ATTGTC	Anneal at 48^{o}C	Novel
EGR1ΔAct-F EGR1ΔAct-R	AGATC**GAATTC**CAGCCTTCGCTAA CCCCT GGTCC**GAATTC**TTAGCAAATTTCA ATTGTC	As standard	Novel
SP6 SP7	TATTTAGGTGACACTATAG TAATACGACTCACTATAGGG	As standard	Promega pGEM^{TM} Vectors
A306 A307	CTATTCGATGATGAAGATACCCCAC CAAACC GTGAACTTGCGGGGTTTTTCAGTAT CTACGAT	Anneal at 68^{o}C	Matchmaker pGADT7 AD Vector (Clontech Laboratories Inc 2009)
A308 A309	TCATCGGAAGAGAGTAGT GTCACTTTAAAATTTGTATACA	As standard	Matchmaker pGBKT7 DNA BD Vector (Clontech Laboratories Inc 2008)

Table 2.4: Oligonucleotides for PCR amplification

2.4.2 Standard PCR

All PCR reactions were set up in 20µl reactions using the following conditions; 10X buffer (1x; Invitrogen), dNTPs (2.5mM; Invitrogen), MgCl2 (2.5mM; Invitrogen), Primers (100ng), Taq Polymerase (2.5 units; Invitrogen) and DNA (100ng).

2.4.3 Nested PCR

To improve the efficiency of amplification of bisulfite treated DNA, the samples under went an additional nested PCR reaction in order to gain more amplification, using 1µl of the original PCR reactions in place of the DNA, using the same conditions as detailed above.

2.4.4 PCR using Bacterial/Yeast plasmids

PCR reactions using either bacterial or yeast plasmids were performed by transferring a colony into the PCR reaction mix in place of DNA using a sterile pipette tip, followed by transfer of the residual colony into a 96 well culture plate containing the relevant media for storage.

2.4.5 Gel Electrophoresis

PCR were products were resolved on agarose gels (1-3%) using routine grade agarose (Biogene, Kimbolton, UK) and 0.5µg/ml of ethidium bromide (BDH, Electran, Poole, England) per 50ml gel in 1x TAE. 5µl of the PCR product was loaded onto the gel with 3µl 1x loading buffer. The 1Kb ladder (Promega) was loaded in a separate well to determine the size of the products. The DNA was electrophoresed at 40-80V for ~40min, and visualised on the BioRad Chemi Doc system using QuantityOne (version 4.4.1) software (BioRad Laboratories Hercules, CA, USA).

2.4.6 Quantitative Real-time PCR

qRT-PCR was performed using the Taqman Gene Expression Assay system (Applied Biosystems) on an ABI PRISM HT7900 Sequence Detection Systme thermal cycler to quantify relative levels of mRNA expression of *EGR1* (Hs00152928 probe mix, Applied Biosystems) and *NAB2* (Hs00195573-m1 probe mix). Expression of *β-actin* (Human ACTB endogenous control probe mix, Applied Biosystems) was used as a reference control.

Each cDNA sample as prepared in 2.2.4 was amplified in triplicate in a 384 well plate (ABgene, Surry, UK) using the following reaction conditions: 2µl of cDNA, 2.5µl sterile H_2O, 5µl Absolute Blue QPCR ROX Mix (ABgene) and 0.5µl of probe mix (20x). A series of dilutions was performed using HCT116 cDNA with the *β-actin* probe mix and the *EGR1* probe mix to generate a standard curve. The PCR reaction conditions were: 50°C – 2min, 95°C – 10min, 95°C – 15secs and 60°c – 1min. The cycle was repeated x40.

The data was analysed using SDS Version 2.3 (Applied Biosystems). In order to determine relative expression levels a standard curve was generated from the dilution series by plotting the threshold cycle (Ct) against the logged quantity of the diluted HCT116 cDNA.

The relative expression of the samples was calculated using linear regression analysis from the standard curve. Only standard curves with a R^2 value of close to 1 were used. The *EGR1* expression values were normalised by dividing by the *β-actin* expression values.

2.5 Sequence analysis

2.5.1 Purification of PCR products

The PCR products were purified using exonuclease 1 (USB, Ohio, US) and shrimp alkaline phosphatase (SAP; USB) in the following reaction mixture: 3.75µl of H_2O, 0.5µl SAP (0.5 U), 0.25µl Exo 1 (2.5 U) and 3µl of PCR product. The mixture was incubated in a Peltier PCT225 thermal cycler (MJ Research) at 37^{o}C for 15min, followed by incubation at 80^{o}C for 15min.

2.5.2 DNA sequencing

Sequencing of the purified DNA was performed in 10µl reactions using ABI PRISM Ready Big Dye Terminator cycle sequencing kit in the following reaction: 100ng of purified DNA, 40ng of primer (forward or reverse), 1µl of Big Dye (Applied Biosystems) and 5µl of H_2O using the following reaction conditions 96^{o}C – 30sec, 50^{o}C – 15sec, 60^{o}C – 4min, x25 cycles.

2.5.3 Precipitation of DNA from sequencing reactions

Following the sequencing reaction the DNA was precipitated by adding 60µl of 95% ethanol and 240µl of NaOAc (0.3M final concentration), and incubating at room temperature for 30min. The samples were spun at 1200rpm for 30min using either Eppendorf Microcentrifuge 5415 R centrifuge (for reactions in eppendorfs) or a Heraeus Multifuge 3 Plus Centrifuge at 2000rpm (for reactions in 96 well plates). The pellets were washed with 70% ethanol and allowed to air dry. The DNA pellets were stored at –20^{o}C. The precipitated reaction products were re-suspended in HiDi^{TM} (Applied Biosystems), heated at 90^{o}C for 2min, and resolved on ABI PRISM® 3100 or 3730 genetic analysers by HGU Technical Services.

2.5.4 Analysis of sequence data

Sequence data was analysed using Consed sequencing program (Gordon, Abajian et al. 1998) and Mutation Surveyor® version 3.30 (Biogene Ltd, UK).

2.6 Cloning and Bacterial culture

All centrifuge steps were done using a Heraeus Multifuge 3 Plus Centrifuge or Thermo Sorvall RC-5B Plus Centrifuge.

2.6.1 Media and Solutions

Luria Broth (L-Broth)*

0.1% weight/volume (w/v) Tryptone (Difco)

0.05% w/v Yeast extract (Difco)

171 mM NaCl

Luria Agar (L-Agar)*

0.1% w/v Tryptone (Difco)

0.05% w/v Yeast extract (Difco)

171 mM NaCl

0.15% w/v Agar (Oxoid Ltd)

Ampicillin Stock Solution

50mg/ml ampicillin (Sigma)

Kanamycin Stock Solution

50mg/ml kanamycin (Sigma)

5-Bromo-4-Chloro-3-Indolyl-β-D-galactoside (X-gal)

40mg/ml X-gal (Sigma) in Dimethylformamide (DMF)

2.6.2 Plasmids

The cloning plasmids, pGBKT7-BD and pGADT7-AD, were obtained from Matchmaker Gal4 Two-Hybrid System 3 (Clontech Labroatories Inc). All plasmids were grown in 250ml L-Broth, with 20mg/ml ampicillin or 10mg/ml kanamycin, overnight with shaking at 37^{o}C. The plasmids were extracted and purified using Maxi Prep Kit (Invitrogen or Qiagen) as per manufacture's instructions. The plasmids were stored as glycerol stocks at –80^{o}C, via addition of 1ml of overnight culture to 300µl sterile glycerol.

2.6.3 Cloning of *EGR1* into pGEM-T® cloning vector

Full length *EGR1* (*EGR1Fl*) and *EGR1* minus its transactivation domain (*EGR1ΔAct*) were amplified from CRC cell line cDNA and the PCR products were purified using a PCR Purification Kit (Qiagen). *EGR1Fl* and *EGR1ΔAct* PCR products were cloned into the pGEM-T® cloning vector (Promega) as per the manufacturer's instructions, and transformed into 50μl of TOP10 Chemically Competent *E.coli* cells (Invitrogen). The transformed cells were spread onto selective L-agar plates containing ampicillin (50μg/μl) and X-gal (40μg/ml). The plates were incubated at 37^{o}C overnight and white colonies were selected and amplified using Sp6 and T7 primers. A colony that was successfully amplified was grown in 250ml L-Broth, with 20mg/ml ampicillin, overnight with shaking at 37^{o}C. Plasmids containing the correct inserts were extracted as demonstrated by plasmid sequencing and purified using a Maxi Prep Kit (Qiagen) as per manufacturer's instructions.

2.6.4 Cloning of *EGR1* into pGBKT7

The pGBKT7 vector and the constructs containing *EGR1Fl* and *EGR1ΔAct* in pGEM-T were cut using EcoR1 in a 100μl reaction at 37^{o}C for 2 hours. Calf Intestine Alkaline Phosphatase (CIAP; Roche) was added to the reactions to prevent self re-ligation. The restriction digest reactions were run out on an agarose gel (0.8-1.5%), with the correct size band extracted and the DNA purified using a Gel Extraction Kit (Qiagen). The ligation reaction was set up using 1μg of cut pGBKT7 vector with 3x cut insert in a 20μl reaction. The reaction was incubated at room temperature for 2 hours. The ligation reaction (5μl) was transformed into 50μl of TOP10 Chemically Competent *E.coli* cell (Invitrogen), plated onto selective L-agar plates containing kanamycin (50μg/μl) and grown overnight at 37^{o}C.

2.6.5 Colony selection and storage

Colonies were selected and amplified using primers A308 and A309, which are present in the pGBKT7 vector, and sequenced to confirm the orientation of the insert. A colony with the correct sequence and orientation was grown in 250ml L-Broth, with 10mg/ml kanamycin, overnight with shaking at 37^{o}C. The EGR1Fl plasmid and the EGR1ΔAct plasmid were extracted and purified using Maxi Prep Kit (Qiagen) as per manufacturer's instructions, and stored as glycerol stocks at -80^{o}C, via 1ml of overnight culture added to 300µl sterile glycerol.

2.7 Yeast Culture

All centrifuge steps were done using a Heraeus Multifuge 3 Plus Centrifuge.

2.7.1 Media and Solutions

YPD broth*

0.01% w/v Yeast extract

0.02% w/v Peptone (Difco)

0.02% glucose

0.003% w/v Adenine

YPD agar*

0.01% w/v Yeast extract

0.02% w/v Peptone (Difco)

0.02% glucose

0.02% w/v Agar (Oxoid Ltd)

0.003% w/v Adenine

SD media*

0.02% w/v Glucose

0.0067% w/v Yeast Nitrogen Base

SD agar*

0.02% w/v Glucose

0.0067% w/v Yeast Nitrogen Base

0.02% Microagar

10x dropout solution

300mg/L Isoleucine (Sigma)

1500mg/L Valine(Sigma)

200mg/L Arginine HCl (Sigma)

300mg/L Lysine HCl (Sigma)

200mg/L Methionine (Sigma)

500mg/L Phenylalanine (Sigma)

2000mg/L Threonine (Sigma)

300mg/L Tyrosine (Sigma)

200mg/L Uracil (Sigma)

Additives

100x Histidine - 200mg Histidine (Sigma) / 100ml H_20

100x Adenine - 200mg Adenine (Sigma) / 100ml H_20

100x Leucine - 1000mg Leucine (Sigma) / 100ml H_20

100x Tryptophan - 200mg Tryptophan (Sigma) / 100ml H_20

10x TE

1M Tris

0.5M EDTA

pH 7.5

10x LiAc

1M LiAc

pH 7.5 with acetic acid

PEG/LiAc

40% PEG 4000

1x TE buffer

1x LiAc

Z buffer

60mM $Na_2HPO_4 \cdot 7H_2O$

40mM $NaH_2PO_4 \cdot H_2O$

10mM KCl

0.1mM $MgSO_4 \cdot 7H_2O$

Z buffer/X-gal solution

100ml Z buffer

0.27ml β-mercaptoethanol

1.67ml X-gal stock solution

Lyticase Solution (Sigma)

5 units/μl in TE buffer

2.7.2 Yeast Culture conditions

The yeast strains Y187 and AH106 were grown in YPD medium or on YPD agar plates. SD media and agar with appropriate nutritional additives (10x dropout solution; 10x histidine; 10x adenine; 10x leucine; 10x tryptophan) was used to grow the yeast strains to be transformed with the bait or library. All yeast were incubated at 30^{o}C and grown for 3-5 days.

2.7.3 Yeast Transformation

The bait constructs in pGBKT7 were transformed into the yeast strain Y187 using a LiAc transformation procedure. The constructs were grown overnight in YPD media at 30^{o}C with shaking. 300ml of YPD was inoculated with 30ml of this starter culture and grown until it reached an OD of 0.4-0.5 using optical densitometry on GeneQuant Pro RNA/DNA calculator UV spectrometer (Amersham Pharmacia biotech). The cells were harvested by centrifugation at 1000xg for 5min, resuspended in 50ml H_2O, spun again at 1000xg for 5 min and resuspended in 1.5ml sterile 1 x TE/1 x LiAc solution.

A solution containing 0.1μg plasmid, 100μg salmon sperm, 100μl of Y187 yeast cells and 0.6ml sterile PEG/LiAc was vortexed for 10sec. The cells were incubated at 30^{o}C for 30min, after which 70μl of DMSO was added and them placed in 42^{o}C water bath for 15min to heat shock. The cells were chilled on ice for 2 min, centrifuged briefly at 14,000 rpm and resuspended in 0.5ml sterile 1x TE. A 1/10 dilution of the cells (100μl) was plated on SD/-Trp (SD media with 10x dropout solution, 10x histidine, 10x adenine and 10x tryptohpan) plates and incubated at 30^{o}C for 3 days.

2.7.4 Yeast Mating

50ml of SD/-Trp (SD media with 10x dropout solution, 10x histidine, 10x adenine and 10x tryptohpan) media was inoculated with the transformed constructs and grown overnight at 30°C with shaking. The cultures were grown until they reached an OD of 0.7-1.0. When the cells were at a concentration of $4x10^8$/ml they were mixed with 250µl of SW480 library cells in AH109 ($2x10^8$/ml) and plated out onto 2 YPDA plates and incubated overnight at 30°C. 5ml of YPDA media was added to each plate and the cells were scraped off, centrifuged at 1000xg for 5 min and resuspended in 2.5ml of SD/-Leu/-Trp (SD media with 10x dropout solution, 10x histidine and 10x adenine) media. Dilutions were made and plated out onto SD/-His/-Leu/-Trp (SD agar with 10x dropout solution with 10x adenine) and SD/-His/-Ade/-Leu/-Trp plates (SD agar with 10x dropout solution). The cells were grown for 3-5 days at 30°C. Colonies were picked from the SD/-His/-Leu/-Trp and the SD/-His/-Ade/-Leu/-Trp plates and inoculated into 96 well plates containing YPD + 25% glycerol and frozen at –80°C.

2.7.5 Isolation of plasmids from yeast

The plasmids were isolated from the yeast cells by growing each interacting colony on a SD/-Leu/-Trp for 3-4 days at 30°C. The colony cells were scraped into an eppendorf tube containing 50µl of TE. Lyticase (10µl) was added to each tube, and the cells were resuspended by vortexing. The samples were frozen at –20°C, thawed, and vortexed again to ensure complete cell lysis.

The plasmid DNA was purified using a phenol;chloroform extraction and ethanol precipitation. TE buffer w as added to each eppendorf to bring the volume to 200µl. 200µl of phenol:chloroform;isoamyl alchol was added, and the sample was vortexed for 5 min. The eppendorfs were centrifigued at 14,000 rpm for 10min. The aqueous (upper) phase was transfered to a fresh tube. 10 M ammonium acetate (8µl) and 95% ethanol (500µl) was added, and frozen at –70°C for 1hr. The samples were

centrifuged at 14,000rpm for 10min, the supernatant was discarded and the dry pellet was resuspended in 20μl of H_2O.

The plasmid DNA was transformed into DH5α cells (Invitrogen) as per the manufacturer's instructions and the prey pGBADT7 plasmid was selected for by plating the transformed cells onto selective L-agar plates containing ampicillin (50μg/μl). Colonies were selected and amplified using A306 and A307 primers, with colonies containing the correct insert extracted and amplified using a mini prep kit (Qiagen) as per manufacturer's instructions.

2.8 Protein Protocols

All centrifugation steps were done on an Eppendorf Microcentrifuge 5415 R centrifuge.

2.8.1 Materials

Normal Lysis Buffer

50mM NaCl

10mM HEPES

500mM sucrose

1mM EDTA

0.5mM spermidine

0.15mM spermine

0.2% w/v Triton X-100

Hypotonic Lysis buffer

350mM NaCl

10mM HEPES

25% w/v Glycerol

0.1mM EDTA

0.5mM spermidine

0.15mM spermine

6x Sample Buffer

20% w/v Glycerol

2% w/v Sodium dodecyl sulfate (SDS)

0.25% w/v Bromophenol blue

1x Stacking buffer

5% w/v β-mercaptoethanol

4x Resolving Buffer

1.5M Tris

0.4% w/v SDS

pH 8.8

4x Stacking Buffer

500mM Tris

0.4% w/v SDS

pH 6.8

10x Running Buffer

250mM Tris

2M Glycine

1% w/v SDS

Semi-Dry Transfer buffer

47mM Tris

40mM Glycine

0.037% w/v SDS

100mM Methanol

10% Resolving Gel

1x resolving Buffer

10% w/v Acrylamide

0.15% w/v Ammonium persulfphate (APS)

0.01% w/v N, N, N', N', tetramethyl-1-2-diaminomethane (TMED)

4% Stacking Gel

1x stacking buffer

4% w/v Acrylamide

0.15% w/v APS

0.01% w/v TMED

IP buffer

300mM NaCL

1% Triton X-100

50mM Hepes (pH 7.4)

2.8.2 Transfection of cell lines

HRT18 cells and SW480 were plated equally into a 6 well plate and grown until ~50% confluent. siRNA transfections were performed using Lipofectamine 2000 (Invitrogen) in Optimem-1 medium (Gibco) and the Stealth^{TM} siRNA (Invitrogen) as per manufacturer's instructions with primers described in Table 2.4. Cells were transfected in antibiotic-free media for 24-48 hours. Cells were harvested as detailed in 2.1.2, the protein was extracted and resolved via Western Blot procedure as detailed below.

Stealth^{TM} siRNA Primers	
EGR1HSS103117	UCUCCCAGGACAAUUGAAAUUUGGU
	AGCAAAUUUCAAUUGUCCUGGGAGA
EGR1HSS103118	GAUCUCUGACCCGUUCGGAUCCUUU
	AAAGAAUCCGAACGGGUCAGAGAUC
EGR1HSS103119	CCAUGGACAACUACCCUAAGCUGGA
	UCCAGCUUAGGGUAGUUGUCCAUGG

Table 2.5: EGR1 siRNA primers

2.8.3 Preparation of protein extracts

Cells were grown until ~80% confluent in a T25cm^2 flask. The cells were harvested, washed in PBS and pelleted by centrifugation at 12,000 rpm for 5min.

2.8.3.1 Preparation of Whole Cell Extracts

To prepare whole cell extracts the cells were resuspended in 200µl of 1X Whole Cell Lysis Buffer (Cell Signalling Technologies) supplemented with the following protease inhibitors; 40µl/ml complete protease inhibitor solution (Roche, Germany), 10µl/ml Pefabloc solution (Roche) and 1µl/ml Pepstatin solution (Roche). After incubation on ice for 30min the cells were centrifuged at 13,000 rpm for 5min at 4^{o}C, and the lysate was transferred to fresh pre-chilled eppendorf tubes.

2.8.3.2 Preparation of Cytoplasmic and Nuclear Extracts:

To prepare cytoplasmic and nuclear extracts the cells were resuspended in 100µl of Normal Lysis Buffer, supplemented with the same protease inhibitors as above. Following an incubation of 5min on ice, the cells were centrifuged at 13,000 rpm for 20sec. The supernatant, the cytoplasmic extract, was removed and transferred to pre-chilled eppendorf tubes. The cell pellet was then resuspended in 30µl Hypotonic Lysis Buffer, supplemented with protease inhibitors as above, and incubated on ice for 30min. The cells were centrifuged at 13,000 rpm for 5min and the supernatant, consisting of the nuclear extract, was removed and transferred to pre-chilled eppendorfs.

2.8.3.3 Preparation of cellular compartments

In order to separate the extracts into four cellular compartments (cytoplamsic, membrane/organelle, nuclear and cytoskeleton) the cells were grown until confluent in a T25cm^2 flask. The cells were then extracted using the ProteoExtract Subcellular Proteome Extraction Kit (Calbiochem) as per the manufacturer's instructions.

All protein extracts were stored at –80^{o}C. The protein concentration of all extracts was determined using a Bradford protein assay, using BSA as a standard control. All samples were measured in triplicate and concentrations calculated against a BSA standard curve.

2.8.4 Western Blot analysis

Protein samples were prepared by dilution with PBS to a suitable protein concentration (30-50ng) and boiled in 3µl of SDS reducing buffer for 5min. The protein samples were separated on a 8-10% resolving acrylamide gel with a 4% stacking gel in 1X SDS running buffer, using 10µl of Kaleidoscope Protein Standard (BioRad) as a marker. The proteins were then transferred onto a PVDF membrane (BioRad) at 10V for 30min, followed by blocking in 5% non-fat milk in 0.15% Tween-10/PBS for 1 hour at room temperature.

The membranes were probed with a primary antibody in milk Tween/PBS at room temperature for 1 hour as standard, unless indicated otherwise (Table 2.5). The membranes were washed in PBS/Tween, 3 times for at least 15 minutes, with shaking, unless stated otherwise. Membranes were incubated in appropriate species of secondary antibodies for 1 hour at room temperature. The membranes were washed in PBS/Tween 3x for 15 minutes with shaking. The protein bands were visualised by chemiluminescence using Luminol reagent (Santa Cruz Biotechnology) by placing the Luminol reagent on the membrane for 1 minute following by placing the membrane in a plastic cover in a film cassette and exposing it to film (Biorad).

Antibody	Manufacturer	Dilution	Conditions
Primary Antibodies			
EGR1 Goat mAB	R&D Systems (AF2818)	1:1000	4°C - overnight
EGR1 (588) Rabbit polyclonal	Santa Cruz Biotechnology (Sc-110)	1:1000	Standard
EGR1 (C-19) Rabbit polyclonal	Santa Cruz Biotechnology (Sc-189)	1:1000	Standard
NOD2 Mouse mAB	Affinity BioReagents (MA1-16611)	1:200	Standard
γ-tubulin Mouse mAB	Sigma Chemical (T6557)	1:5000	Standard
α-tubulin Mouse mAB	Sigma Chemical (T9026)	1:1000	Standard
HP1-alpha Rabbit mAb	Cell Signaling Technology (C7F11)	1:1000	Standard
B23-(3F291) (NPM1)	Santa Cruz Biotechnology (Sc-70392)	1:1000	Standard
eEF1α	Upstate (Millipore) (05-235)	1:1000	Standard
eEF1A1	Novel- gift from Cathy Abbott	1:1000	4°C - overnight
eEF1A2	Novel- gift from Cathy Abbott	1:1000	4°C - overnight
β-actin mAB Mouse	Sigma Chemical (A1978)	1:160,000	Standard
Secondary Antibodies			
Goat anti-Mouse IgG-HRP	Santa Cruz Biotechnology (Sc-2005)	1:2000	Standard
Donkey anti-Goat IgG-HRP	Santa Cruz Biotechnology (Sc-2056)	1:2500	Standard
Goat Anti-Rabbit IgG - Horseradish Peroxidase	Amersham Biosciences (RPN4301)	1:2000	Standard

Sheep Anti-Mouse IgG - Horseradish Peroxidase	Amersham Biosciences (RPN4201)	1:2000	Standard

Table 2.6: Antibodies for Western Blot analysis

2.8.5 Immunoprecipitation

SW480 cells were seeded into a T25cm^2 flask and grown until 80-90% confluent. The cells were harvested and the protein was extracted in 200µl of lysis IP buffer with protease inhibitors added and the protein concentration was determined by Bradford assay. The protein extracts were pre-cleared in washed (IP buffer) agarose G beads for 1 hour at 4^{o}C and 500µg of protein lysate was incubated in 2µg of antibody overnight at 4^{o}C while rotating. 15µl of pre-washed agarose G beads was added and incubated at 4^{o}C for 1 hour while rotating. The beads were washed and 50µl of 2x running buffer was added. The samples were boiled for 5 min at 95^{o}C, and centrifuged at 14000rpm at room temperature for 5min using an Eppendorf Microcentrifuge 5415 R. The supernatant was separated on a 10% resolving gel and visualised by Western Blot analysis.

2.9 Immunofluorescence

SW480 and HRT18 cells were seeded into a 6 well plate containing a sterile coverslip and grown overnight at 37^{o}C. The media was removed and cells washed in 5ml cold PBS. The cells were fixed with 1ml of 4% formaldehyde and incubated at room temperature for 20min. After removal of the formaldehyde, the cells were rinsed several times with PBS, and a solution of 0.5% Triton X-100 was added to the cells for 2min at room temperature. The cells were washed with PBS x3 for 5min while shaking the cells and then blocked with 5% BSA in PBS all day.

The cells were then washed with PBS x3 for 5min as before and incubated with a primary antibody in 5% BSA in PBS overnight at 4^{o}C (Table 2.6). After washing with PBS x3 for 5min, the cells were incubated with a secondary antibody in 5% BSA for 1hr at room temperature in the dark. The cells were washed x3 for 10min as before and the coverslips were placed onto microscope slides containing a few drops of DAPI-Vectashield® (Vector Laboratories Inc., USA), containing 1μg/ml 4',6-diamidino-2-phenylindole (DAPI). The slides were stored in the dark at 4^{o}C and the images were visualised using fluorsecence microscopy.

Antibody	Manufacturer	Dilution used
Primary Antibodies		
EGR1 Goat mAB	R&D Systems AF2818	1:50
NOD2 Mouse mAB	Affinity BioReagents MA1-16611	1:200
γ-tubulin Mouse mAB	Sigma Chemical T6557	1:200
α-tubulin Mouse mAB	Sigma Chemical T9026	1:200
Secondary Antibodies		
Texas Red-AffiniPure Bovine Anti-Goat IgG	Jackson Immunoresearch Laboratories, 805-075-180	1:200
Texas Red-AffiniPure Fab Frgament Donkey Anti-Mouse IgG	Jackson Immunoresearch Laboratories, 715-097-003	1:200

Table 2.7: Antibodies for immunocytochemistry analysis

3 Chapter 3

3.1 Introduction

The expression of *EGR1* has been found to be dysregulated in numerous types of cancers including breast carcinoma, glioblastoma, osteogenic sarcoma, fibrosarcoma, and esophageal carcinoma (Huang, Liu et al. 1995; Huang, Fan et al. 1997; Calogero, Arcella et al. 2001; Wu, Chen et al. 2001). In contrast, *EGR1* expression was found to be increased in gastric cancer compared with normal mucosa (Kobayashi, Yamada et al. 2002) and *EGR1* is known to be over-expressed in prostate cancer, with differential expression observed between prostate epithelial and stromal tissue (Svaren, Ehrig et al. 2000; Gregg, Brown et al. 2010).

In silico studies in the Colon Cancer Genetics Group (CCGG) (University of Edinburgh) identified that *EGR1* is downregulated in the tumour mucosa of patient with CRC compared with normal colonic mucosa, indicating a role in carcinogenesis (James Prendergast/Susan Farrginton pers comm).

A microarry study by Hong et al (2007) describes an up-regulation of *EGR1* expression in early onset CRC patients, where they observed higher *EGR1* expression levels in the normal mucosa of 21 CRC patients compared with mucosa from 10 healthy controls indicating that EGR1 expression is up-regulated in the normal mucosa of colorectal cancer patients. They suggest that as they observed differential gene expression in seven genes between the mucosa from patients and controls, that although the normal mucosa appears morphologically normal in patients it is "already primed for carcinogenesis" (Hong, Ho et al. 2007). Given this finding an analysis was conducted using Oncomine (Rhodes, Yu et al. 2004) in order to compare the expression of EGR1 in normal and tumour mucosa in published microarray data. This analysis showed that there is a significant difference in EGR1 expression, with a 7.539 fold higher expression of EGR1 in colon tissue (n=12) compared with colorectal carcinoma (n=70) (Hong, Downey et al. 2010). A smaller

study saw significantly higher EGR1 expression in colon (n=10) compared with carcinoma with a 3.459 fold increase, and a 3.428 fold higher expression from colon (n=10) to adenoma (n=5) (Skrzypczak, Goryca et al. 2010). However several studies showed lower expression of EGR1 in normal colon compared with colorectal adenocarcinoma. The most significant result was the study by Skrzypczak, which showed that there was significantly lower expression in colorectal tissue (n=24) compared with colorectal adenocarcinoma, with a 3.373 fold difference. This study however showed no significant difference in the EGR1 expression in normal and colorectal carcinoma. Three further studies showed a significant difference in EGR1 expression from colon to colorectal adenocarcinoma. This analysis in oncomine showed that there is differential expression of EGR1 in normal tissue and colorectal tissue, with an indication that EGR1 expression is lower in colorectal adenomas compared to normal, but that expression may be gained in invasive carcinomas.

It is known that patients with IBD are at an increased risk of developing CRC, and this risk further increases with duration of symptoms and severity of inflammation and dysplasia. Although IBD associated cancer only accounts for 1-2% of overall CRC cases, it accounts for 10-15% of all deaths in patients with IBD (Munkholm 2003). Interestingly, Subbaramaiah et al (2004) observed that in both Crohn's disease and ulcerative colitis patients, there was an increase in the levels of EGR1 protein in the inflamed mucosa compared with non-inflamed (Subbaramaiah, Yoshimatsu et al. 2004). Experiments conducted by the Gastrointestinal unit (GI) (University of Edinburgh), examined the expression of *EGR1* in colonic biopsies and also showed that *EGR1* is differentially regulated between un-inflamed and inflamed mucosa. In collaboration a large scale case-control study was conducted by both the CCGG and the GI unit which demonstrated an association between common variants of *EGR1* with disease phenotype in both CRC and IBD.

Therefore it is of interest to determine why *EGR1* may be differentially regulated in colorectal disease and to what extent this differential regulation occurs in the Scottish population. The aim of this chapter is to examine the differential expression of *EGR1* in colonic mucosa of healthy controls, uninflamed mucosa of patients with IBD and from matched normal mucosa and tumour tissue of colorectal cancer patients. We

investigated whether genetic variation correlated with altered gene expression by looking at the common EGR1 variants identified. Given that EGR1 has a large CpG promoter region that may undergo methylation, the methylation status of the EGR1 promoter was examined. EGR1 expression is regulated via the MAPK signalling pathway and given the importance of RAF and RAS mutations in CRC the somatic tumour mutation status of RAF and RAS were also investigated in patients with CRC.

3.2 Methods

3.2.1 Quantitative RT-PCR

In order to determine the relative expression of *EGR1* qRT-PCR was performed. In a single experiment each assay was conducted in triplicate, with each experiment conducted three times. A standard curve was generated for β-actin and *EGR1* in order to estimate relative expression levels, using linear regression analysis. All the standard curves used to calculate the relative expression levels had a R^2 value close to 1 (>0.92). The *EGR1* expression values were normalised by dividing by the relevant β-actin expression values (as described in ABI PRISM 7700 Sequence Detection System publications).

3.2.2 CRC patients

Colonic biopsies from patients with colorectal cancer (CRC) were available for *EGR1* expression analysis, from matching normal mucosa and tumour. The RNA and DNA was extracted using TRIzol (by Rebecca Barnetson), and the expression of EGR1 was analysed using qRT-PCR.

3.2.3 IBD patients

Colonic biopsies from patients with ulcerative colitis (UC) and Crohn's disease (CD) were available for *EGR1* and *NAB2* expression analysis. All biopsies were taken from un-inflamed tissue. The samples were treated with various inflammatory agents, LPS, MDP, PGN and TNF, at 1μg/ml for 24hrs, to stimulate an inflammatory response. Biopsies were also available from 'healthy' controls (HC) who are healthy individuals undergoing screening for a family history for CRC. The IBD biopsy samples were prepared by the GI unit, using an ex vivo tissue culture protocol (Chapter 2.1.4.2).

3.2.4 Statistical analysis

All of the data was tested for normality using the Shapiro-Wilk test. In the case of the CRC data, as it did not meet the conditions for normality a log transformation was used. The Student paired t-test was used in order to determine differences in *EGR1* expression between the matched normal and tumour samples. For the IBD data, which again did not meet the conditions of normality several different methods of analyzing the data were discussed with quantitive statistical geneticists (Albert Tenesa, Stephane Ballereau, Andy Sims and Jennifer Huffman). Initially the data was analysed using ANOVA but as ANOVA requires parametric data, the non-parametric Kruskal-Wallis rank sum test was then considered. However this model does not allow for an unbalanced data set, which excluded a lot of the data. It also did not account for the repeated measures in the experimental design. Finally a mixed model ANOVA was decided on, which would account for the repeated measures in the data and allow an unbalanced data set. Several transformations were tested on the data, including log and square-root, however only a rank transformation transformed the data sufficiently to a normal distribution. The IBD patient data was therefore transformed using a rank transformation and analysed using a mixed effect ANOVA. All of the data analysis was done using 'R' (R Development Core Team 2009), using the following packages; 'lme4', 'nlme' and 'multcomp' (Hothorn, Bretz et al. 2008; Pinheiro, Bates et al. 2009). Individual patient analysis was conducted using the Wilcoxon rank sum test, comparing each treatment against the untreated value, on non transformed data. Data was considered to be significant at a p value <0.05, which is denoted by a * symbol, a p value of <0.01 is denoted by **, and $p=<0.001$ denoted by ***.

3.2.5 Genotype analysis

All of the CRC and IBD patients were genotyped for three *EGR1* variants. DNA was extracted from the biopsies (as in 2.2 and 2.3) and amplified by PCR. Variants in the three SNPs (rs3813221, rs11734810 and rs11748488) were identified using the sequencing software Mutation Surveyor and polyphredPhrap/Consed.

3.2.6 *KRAS/BRAF* mutation analysis

The *KRAS* and *BRAF* mutation status of the CRC patients was determined. DNA was extracted and amplified by PCR and visualised using Mutation Surveyor. HT29 CRC cell line was used as a previously reported positive control for the BRAF V600E mutation (Davies, Bignell et al. 2002; Seth, Crook et al. 2009).

3.2.7 Bisulfite sequencing

The methylation status of the *EGR1* promoter was investigated using bisulfite sequencing. Changes in the methylation of DNA can occur on cytosine residues, within CpG rich regions of the gene. Bisulfite treatment of DNA causes the cytosine residues to become converted to uracil residues, however cytosines that are methylated remain unchanged. This DNA is then amplified using PCR primers that are specific to unmethylated DNA, the PCR product is then sequenced, where the non-methylated regions displayed as thymines and the methylated regions remaining cytosines (Figure 3.1).

Primers were designed using Methprimer (Li and Dahiya 2002), which covered 80% of the promoter CpG island and allowing sequencing of ~75% of the CpGs, corresponding with the predicted CpG island using UCSC genome browser. A normal mucosa sample was treated with a DNA methylase enzyme, S-(5'-Adenosyl)-L-homocysteine (SAM), as a methylated positive control, which cuases the cytosine residues to undergo methylation.

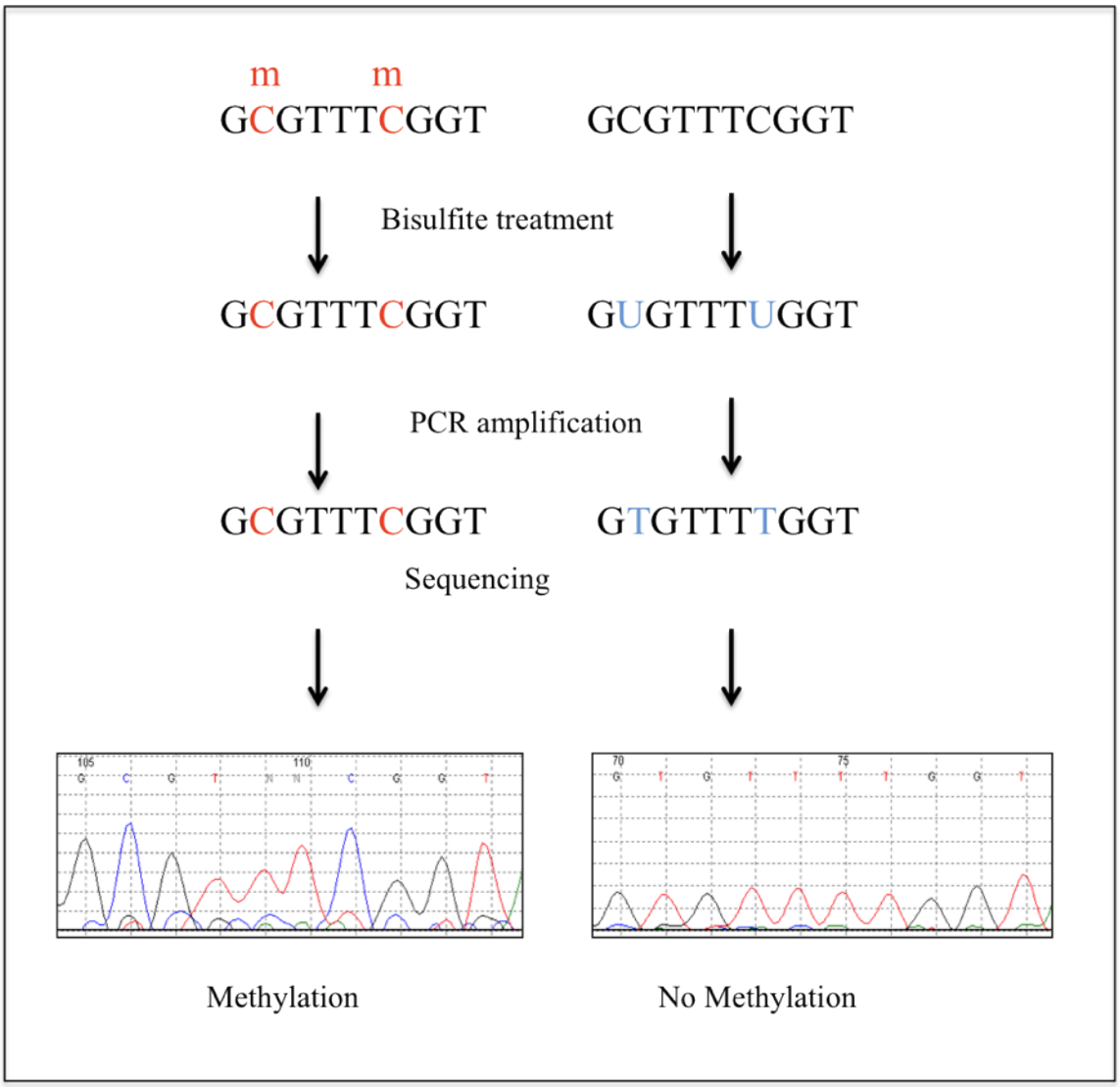


Figure 3.1: Illustration of bisulfite sequencing of DNA

Bisulfite treatment of DNA causes the cytosine residues to become converted to uracil residues (U, denoted in blue). Cytosines that are methylated (mC, denoted in red) remain unchanged. DNA is then amplified using PCR primers and the PCR product is sequenced. Non-methylated CpGs are displayed as thymines (T, denoted in blue) and the methylated CpGs remain as cytosines.

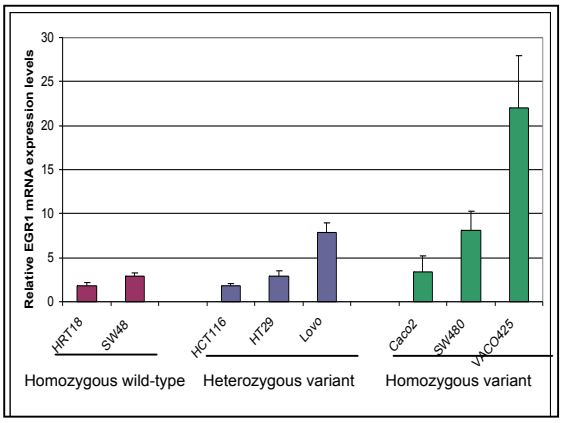
3.3 Results

3.3.1 Relative *EGR1* mRNA expression levels in CRC cell lines and tissue specific expression levels

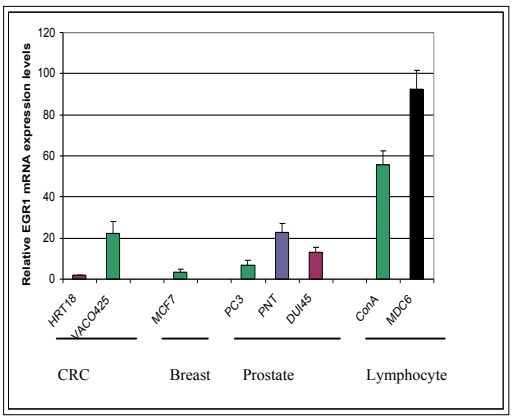
A series of CRC cell lines were initially analysed to determine their relative *EGR1* mRNA expression levels (Figure 3.2 A). There is relatively low *EGR1* expressed in all of the eight colorectal cancer cell lines examined, with the lowest levels evident in HRT18 and HCT116 cells. Vaco425 and SW480 have the highest level of *EGR1*. In order to determine if there is any correlation with the three associated variants for *EGR1*, the CRC cell lines were genotyped for the three SNPs (rs3813321, rs11734810 and rs11748488). SNPs rs3813321 is 856 base-pairs upstream of EGR1, rs11734810 is intronic (1.235 base-pairs from the start of the coding regions of EGR1) and rs11748488 is 579 base-pairs downstream from EGR1. LoVo, HCT116 and HT-29 are heterozygous for the three risk variants. Vaco425 and SW480 are homozygous for all three variants and HRT18 and SW48 are homozygous wild-type.

Other cell lines were examined for *EGR1* expression to assess differences in tissue expression (Figure 3.2 B). There are considerably higher levels of *EGR1* expression in cell lines derived from lymphocytes (MDC6 and ConA), compared to the CRC cell lines. The two prostate cell lines, DUI 145 and PNT show expression levels similar to Vaco425 with the breast cancer cell line MCF7 expressing much lower *EGR1* levels, similar to HRT18 and HCT116. These data agrees with published data where it has been reported that there are higher levels of *EGR1* in prostate cancers, with down-regulation of *EGR1* in breast cancer (Huang, Fan et al. 1997; Svaren, Ehrig et al. 2000). These results show that EGR1 is differentially expressed in different tissues with extremely high levels being observed in the lymphocyte lineages and relatively low levels in cancer tissue of breast, colon and prostate. The assay appears to robustly detect a range of higher and lower levels of *EGR1*.

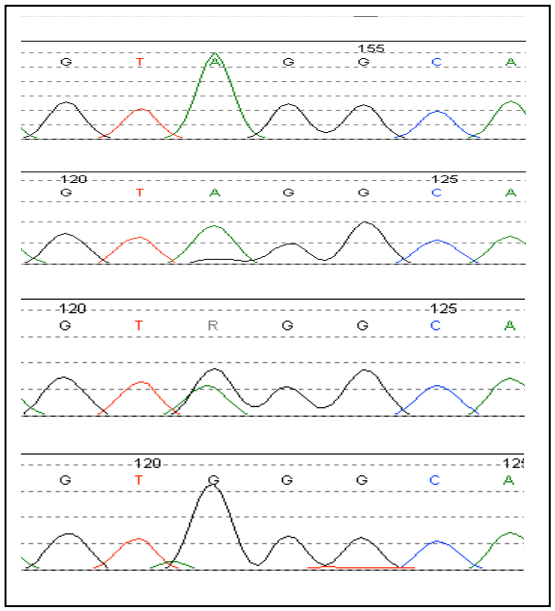
A.



B.



C.



Reference sequence – *EGR1*

HRT18 – Homozygous WT

HCT116 – Heterzygous variant

SW480 – Homozygous variant

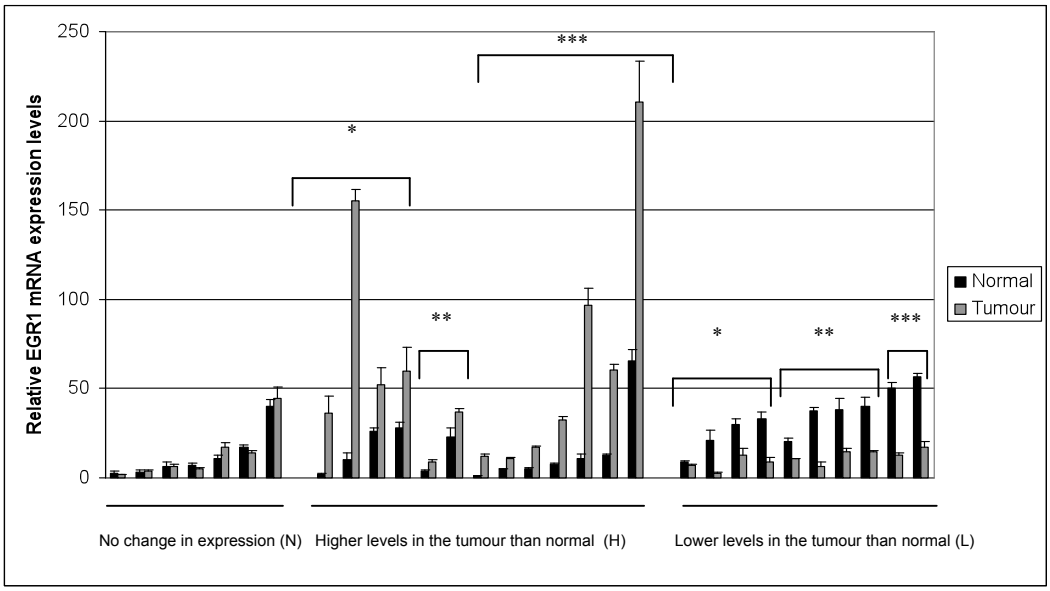
Figure 3.2: Expression and genotyping of EGR1 in cancer and non-cancer cell lines

A. RNA from several different CRC cell lines was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to β-actin. **B.** RNA from several different cell lines was isolated and analysed as above. MCF-7 is a breast adenocarcinoma, DUI145 a prostate carcinoma, PC3 a prostate adenocarcinoma, PNT a normal prostate cell lines, and ConA and MDC6 cell lines are derived from lymphocytes. **C.** Cell line DNA was isolated, amplified by PCR and sequenced to determine the genotype of three *EGR1* variants. The sequencing demonstrates an example of homozygous wild-type, homozygous variant and heterozygous variant of SNP rs11748288.

3.3.2 Relative *EGR1* mRNA expression in CRC patients

The relative *EGR1* mRNA expression was analysed in 30 patients with CRC, using matched normal mucosa and tumour. A student's paired t-test was used to determine if there was any significant difference in the expression of *EGR1* in the matched normal mucosa and tumour. Interestingly there appears to be differential expression of *EGR1* between the normal mucosa and tumour, with a population of the patients showing higher levels of *EGR1* in the tumour (group H) compared with the normal and the other population showing lower levels of *EGR1* in the tumour (group L) (Figure 3.3 A). Of the 30 patients, 13 show a significantly higher level of *EGR1* in the tumour compared with the matched normal using a paired student t-test, 10 show a significantly lower level of *EGR1* in the tumour and 7 samples do not show a significant difference (group N) (Figure 3.3 B).

A.



B.

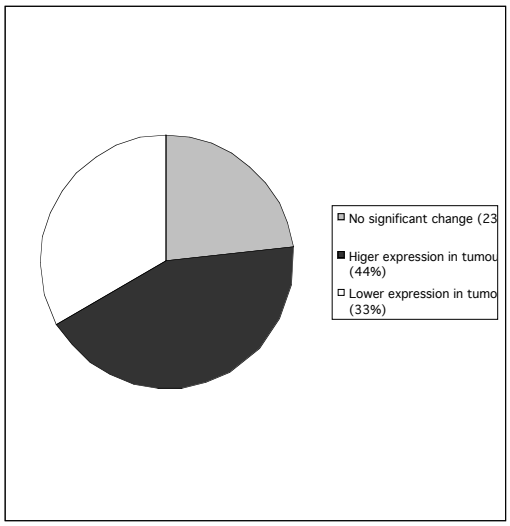


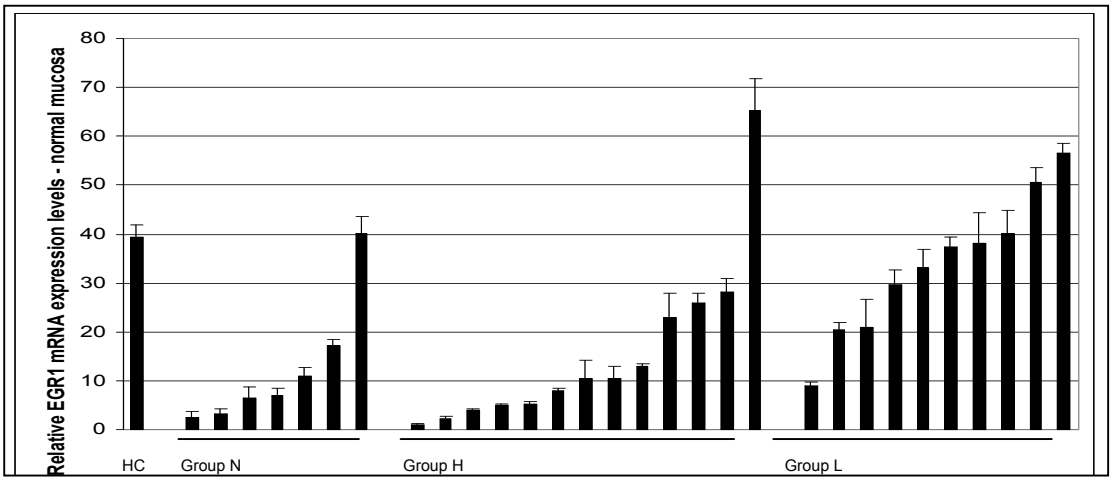
Figure 3.3: Differential expression of *EGR1* in matched normal mucosa and tumour

A. RNA from 30 matched normal and tumour patient samples was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to *β-actin*. The values were transformed using log to normalise the data for statistical analysis. Student's paired t-test was used to determine significant changes, with significance set at < 0.05. **B**. Illustration of the three groups showing differential *EGR1* expression levels in the CRC patient samples between normal and tumour.

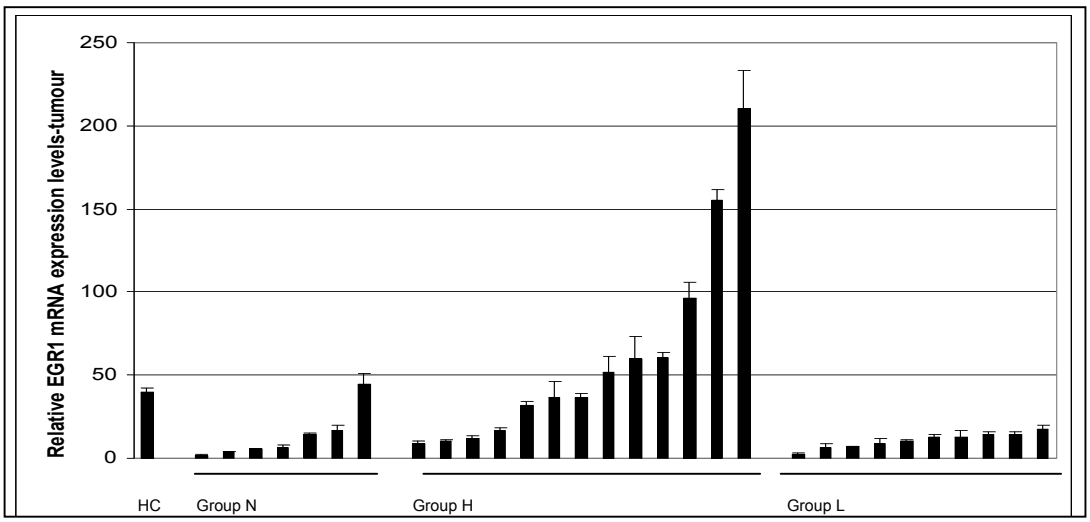
When looking at the expression of *EGR1* in the normal mucosa it varies between the patient samples, ranging from low levels comparable with the cell lines and relatively much higher levels (Figure 3.4 A). It appears that, save for a few outliers, that the normal mucosa levels in those patients in Group H have relatively low *EGR1* levels compared with Group L. Interestingly Group L, the group of patients with higher levels of *EGR1*, have *EGR1* expression levels comparable to the *EGR1* expression in normal healthy mucosa.

Plotting the relative *EGR1* mRNA levels against whether they show higher or lower levels in the matching tumour illustrates that, on average, those patients which show higher levels of *EGR1* in the tumour do have significantly lower levels of *EGR1* compared with the patients that have lower levels in the tumour (Figure 3.4 C). Overall it appears that if there are low levels of *EGR1* in the normal mucosa the corresponding level of *EGR1* in the tumour will be higher and conversely if there are high levels of *EGR1* in the normal mucosa, the tumour will have low levels of *EGR1* (Figure 3.4 D). The patients that show no significance difference between the normal and tumour tended to have low levels of *EGR1* in the normal mucosa.

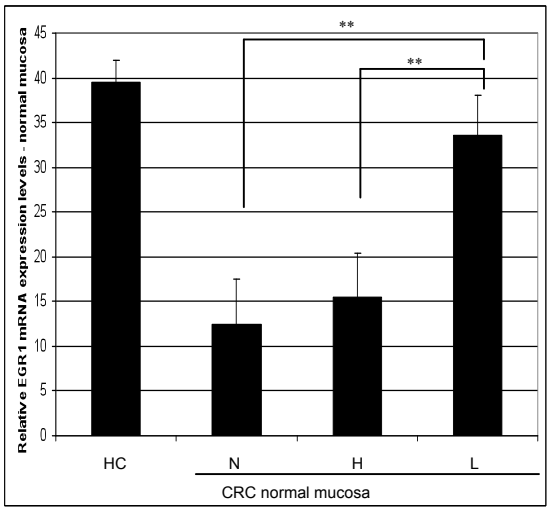
A.


Relative EGR1 mRNA expression levels - normal mucosa
80
70
60
50
40
30
20
10
0
HC
Group N
Group H
Group L

B.


Relative EGR1 mRNA expression levels-tumour
250
200
150
100
50
0
HC
Group N
Group H
Group L

C.


Relative EGR1 mRNA expression levels - normal mucosa
45
40
35
30
25
20
15
10
5
0
**
**
HC
N
H
L
CRC normal mucosa

D.

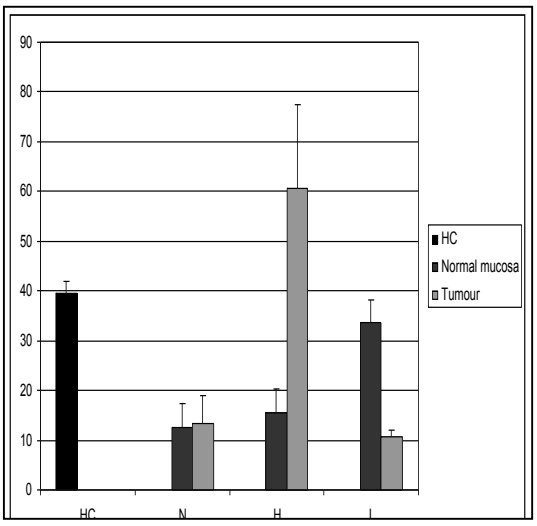

90
80
70
60
50
40
30
20
10
0
HC
Normal mucosa
Tumour
HC
N
H
L

Figure 3.4: Relative *EGR1* mRNA expression levels in normal mucosa

The data was divided based on differences in expression between normal mucosa and tumour. **A.** The relative *EGR1* mRNA expression values of 30 CRC patients normal mucosa was plotted. The average *EGR1* expression of healthy controls (n=17) is plotted (HC) to illustrate the level of *EGR1* expression in healthy mucosa **B.** The relative *EGR1* mRNA expression values of 30 CRC patients tumour was plotted. **C.**The average values of the normal mucosa are shown to determine the *EGR1* expression level of the population. A mixed effect ANOVA model was used to determine the significant difference with the significance set at < 0.05. **D.** The average values of the normal mucosa and tumour in the three populations.

3.3.3 Genotype of *EGR1* variants in matched normal and tumour patient samples

Variation at the genomic level can alter transcriptional activity, hence three common SNPs of EGR1 were genotyped in the matched normal and tumour samples, the results of which are presented in Table 3.1. Three patients are homozygous wild-type of all three SNPs, 12 patients are heterozygous variants for all three SNPs and 5 patients are homozygous variant. Two patients (#9 and #7) show loss of heterozygosity for all three variants in the tumour. There are 5 patients in which show recombination in both the normal mucosa and tumour, for example patient 5 is heterozygous variants for rs3813221 but homozygous variant for both rs11743810 and rs11748288.

No change in expression

	rs3813221	rs11743810	rs11748288
20N	CC	CC	AA
20T	CC	CC	AA
4N	CT	CT	AG
4T	CT	CT	AG
8N	CT	CT	AG
8T	CT	CT	AG
6N	CT	CT	AG
6T	CT	CT	AG
3N	TT	TT	GG
3T	TT	TT	GG
23N	TT	TT	GG
23T	TT	TT	GG
5N	CT	TT	GG
5T	CT	TT	GG

Higher EGR1 expression in tumour

	rs3813221	rs11743810	rs11748288
17N	CC	CC	AA
17T	CC	CC	AA
31N	CC	CC	AA
31T	CC	CC	AA
10N	CT	CT	AG
10T	CT	CT	AG
11N	CT	CT	AG
11T	CT	CT	AG
14N	CT	CT	AG
14T	CT	CT	AG
28N	CT	CT	AG
28T	CT	CT	AG
24N	CT	CT	AG
24T	CT	CT	AG
19N	TT	TT	GG
19T	TT	TT	GG
29N	TT	TT	GG
29T	TT	TT	GG
9N	CT	CT	AG
9T	TT	TT	GG
1N	TT	CC	GG
1T	TT	CC	GG
2N	TT	CC	AA
2T	CC	CC	AA
27N	TT	CT	AG
27T	CT	CT	AG

Lower EGR1 expression in tumour

	rs3813221	rs11743810	rs11748288
18N	CT	CT	AG
18T	CT	CT	AG
22N	CT	CT	AG
22T	CT	CT	AG
25N	CT	CT	AG
25T	CT	CT	GG
30N	CT	CT	AG
30T	CT	CT	AG
12N	TT	TT	GG
12T	TT	TT	GG
7N	CT	CT	AG
7T	TT	CC	GG
15N	CT	TT	GG
15T	CT	TT	GG
16N	CT	TT	GG
16T	CT	TT	GG
26N	CC	CT	AG
26T	CC	CT	AG
21N	TT	TT	GG
21T	CT	TT	GG

Table 3.1: Genotype of *EGR1* variants in CRC patient samples

Figure 3.5 A shows the frequency of the alleles of the three SNPs in the three populations, i.e. those samples that have no change in expression between normal and tumour (Group N), those that have higher levels of *EGR1* in the tumour (Group H) and those that have lower levels of *EGR1* in the tumour (Group L). There does not appear to be any difference in the frequency of the alleles in the three groups, however there are no homozygous wild type alleles present for rs11743810 and rs1748288 in group L, (the group with significantly lower levels of *EGR1* in the tumour and high EGR1 expression levels in the normal mucosa comparable with the health controls, Figure 3.4) suggesting again that perhaps the presence of the *EGR1* variants correlates with differential *EGR1* expression. However when the alleles of the three variants were plotted against the *EGR1* gene expression (Figure 3.5 B) there are no obvious differences in EGR1 gene expression with alleles of the SNPs.

A. B.

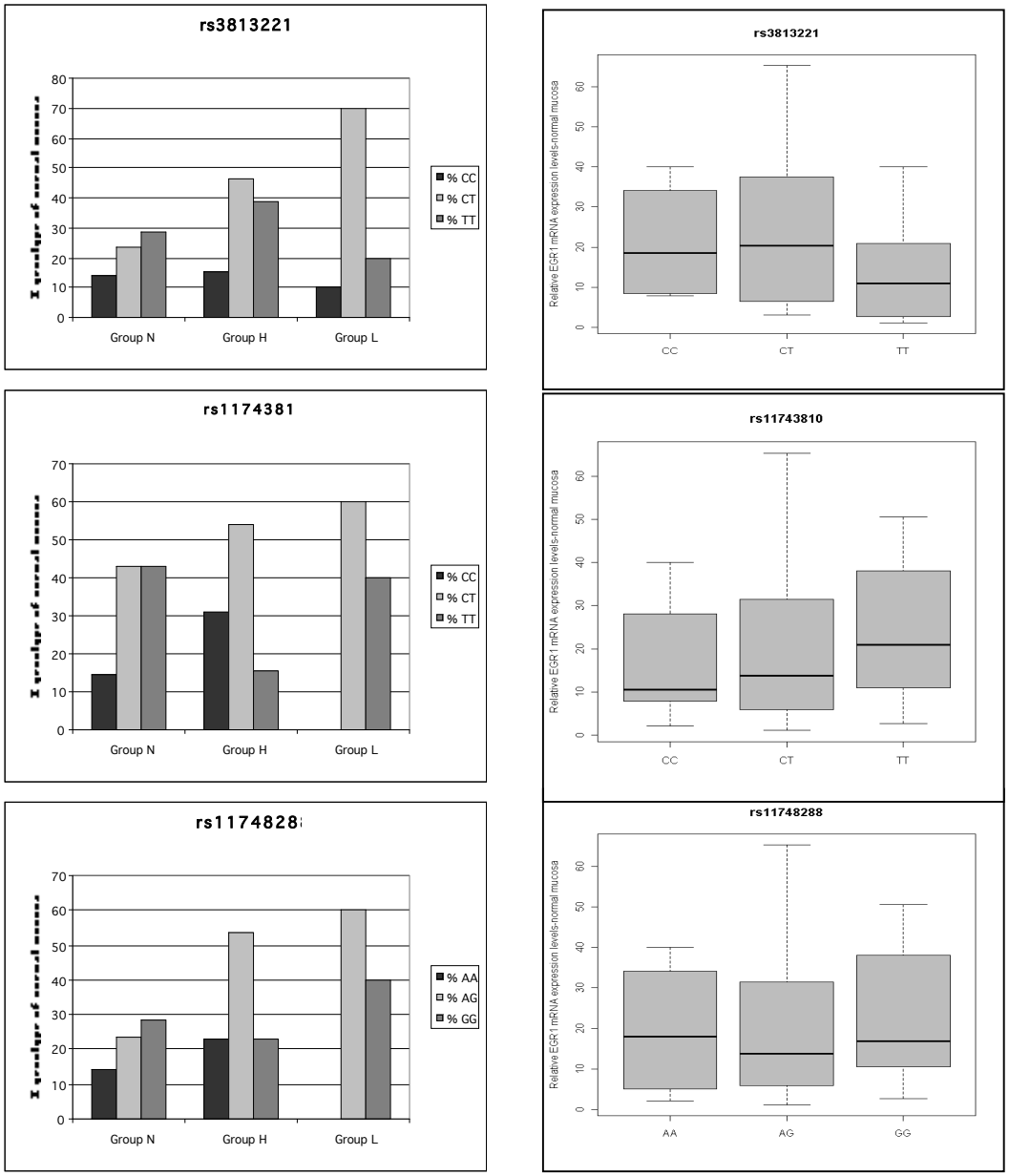


Figure 3.5: Genotype analysis of *EGR1* variants in CRC patients

DNA from matched normal and tumour of 30 CRC patients, was isolated, amplified by PCR and sequenced to determine the genotype of three *EGR1* SNPs, rs3813221, rs11743810 and rs11748288. **A** The data was grouped based on differences in expression between normal mucosa and tumour as before, illustrating what percentage of alleles is present for each SNP in the three groups. **B** The relative *EGR1* expression levels of the normal mucosa are plotted against the alleles of the three SNPs to determine if there is a correlation with gene expression and presence of the *EGR1* variants.

3.3.4 *KRAS* and *BRAF* mutation status of CRC patients

The mutation of two oncogenes, *KRAS* and *BRAF*, are early events in the development of CRC. Both *KRAS* and *BRAF* are members of the MAP kinase pathway and are mutated in ~30-50% and ~10-18% of all CRC respectively (Fearon and Vogelstein 1990; Ilyas, Straub et al. 1999; Davies, Bignell et al. 2002). Activating mutations of *KRAS* most often occur codon 12 and 13 (~90%), although mutation can be present at other codons such as 61 and 146, and result in a constitutively active Ras protein (Nagasaka, Sasamoto et al. 2004; Lee, Cho et al. 2008). Activating *BRAF* mutations are found at exon 11 and 15, and most commonly occur at codon 600 (~80%) resulting in an amino acid change from a valine to a gluatmate. It has been reported that this change greatly increases the kinase activity of BRAF (Davies, Bignell et al. 2002; Nagasaka, Sasamoto et al. 2004; Wan, Garnett et al. 2004).

The somatic tumour mutations status of *KRAS* and *BRAF* of the CRC patients was determined using PCR. Of the 30 patients, nine have a mutation in codon 12 of the KRAS protein. Three types of mutations were seen in codon 12 (Figure 3.6 A, Table 3.2), the most common being a substitution from a GGT -> GAT, resulting in a change from a glycine (G) to an asparatic acid (D). There were three patients with a GGT -> GTT, resulting in a valine (V) and one patient with a GGT -> TGT substitution that results in a cysteine (C). As *BRAF* V600E mutations are relatively uncommon, the CRC cell line HT29 was used as a reported positive control for this mutation (Davies, Bignell et al. 2002; Seth, Crook et al. 2009). Two patients were found to have a V600E mutation in *BRAF* (Figure 3.6 B). There were no *KRAS* or *BRAF* mutation detected in the normal mucosa of any of the CRC patients.

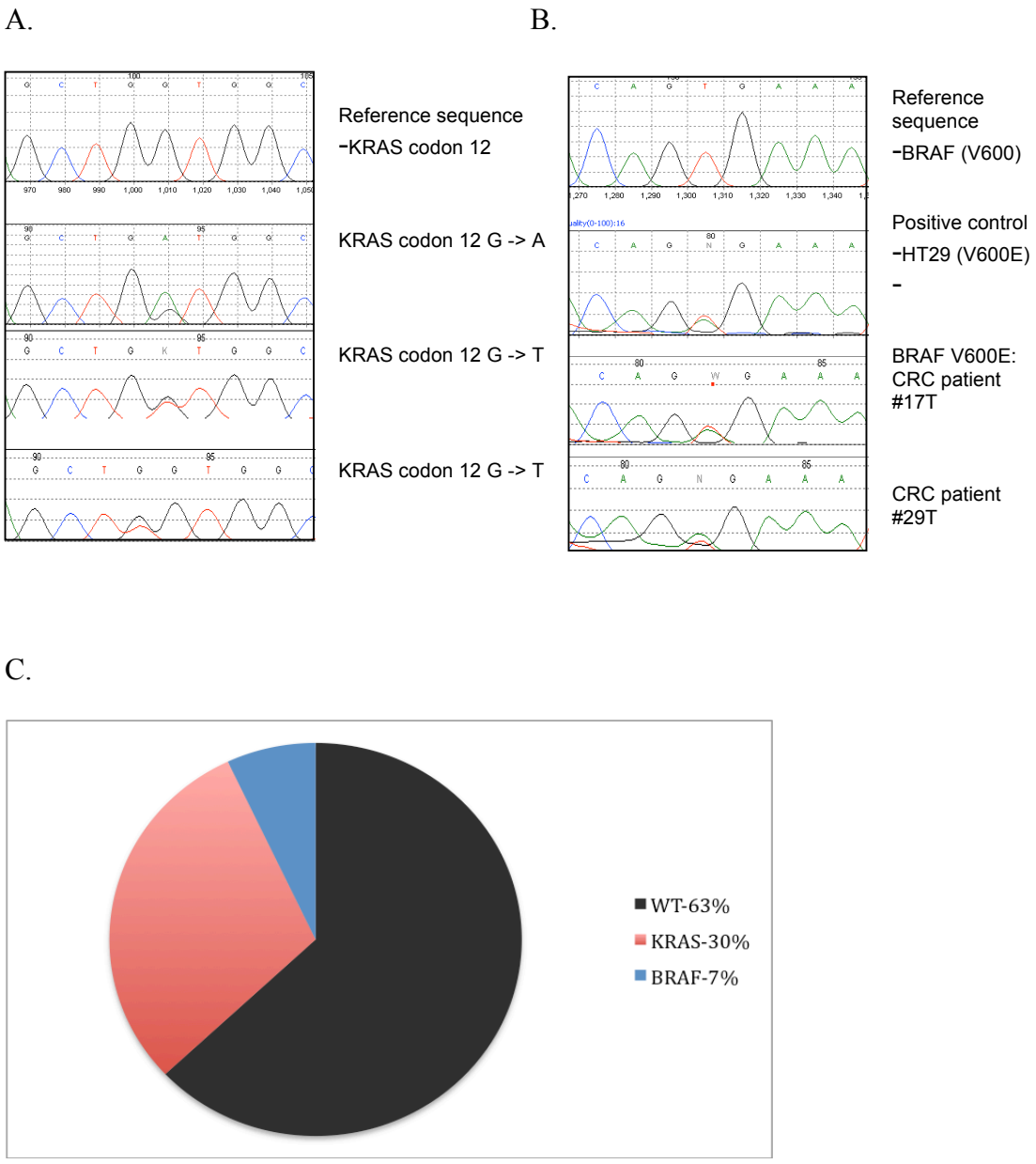


Figure 3.6: *KRAS* and *BRAF* mutation status of CRC patients

DNA from matched normal and tumour of 30 CRC patients was isolated, amplified by PCR and sequenced to determine the mutation status of *KRAS* and *BRAF*. **A.** Sequencing from CRC tumours indicating the three types of *KRAS* mutations that were found in the CRC patients (n=30). All mutations found (n=9) were in exon 12 of the *KRAS* gene. **B.** Sequencing from the two CRC tumours that have a V600E mutation in *BRAF*. HT29 cell line DNA was used a positive control for this mutation. **C.** Illustration of the percentage of *BRAF* and *KRAS* mutations found in tumours of the 30 CRC patients.

Patient	*BRAF* mutation	*KRAS* mutation	cDNA change
IT	wt	G12D-KRAS	GGT -> GAT
2T	wt	wt	
3T	wt	wt	
4T	wt	G12D-KRAS	GGT -> GAT
5T	wt	wt	
6T	wt	G12V-KRAS	GGT -> GTT
7T	wt	wt	
8T	wt	G12V-KRAS	GGT -> GTT
9T	wt	G12D-KRAS	GGT -> GAT
10T	wt	wt	
11T	wt	wt	
12T	wt	wt	
13T	wt	wt	
14T	wt	wt	
15T	wt	wt	
16T	wt	wt	
17T	V600E	wt	GTG -> GAG
18T	wt	wt	
19T	wt	G12V-KRAS	GGT -> GTT
20T	wt	G12D-KRAS	GGT -> GAT
21T	wt	wt	
22T	wt	wt	
23T	wt	wt	
24T	wt	G12D-KRAS	GGT -> GAT
25T	wt	wt	
26T	wt	G12A-KRAS	GGT -> TGT
27T	wt	wt	
28T	wt	wt	
29T	V600E	wt	GTG -> GAG
30T	wt	wt	
31T	wt	wt	

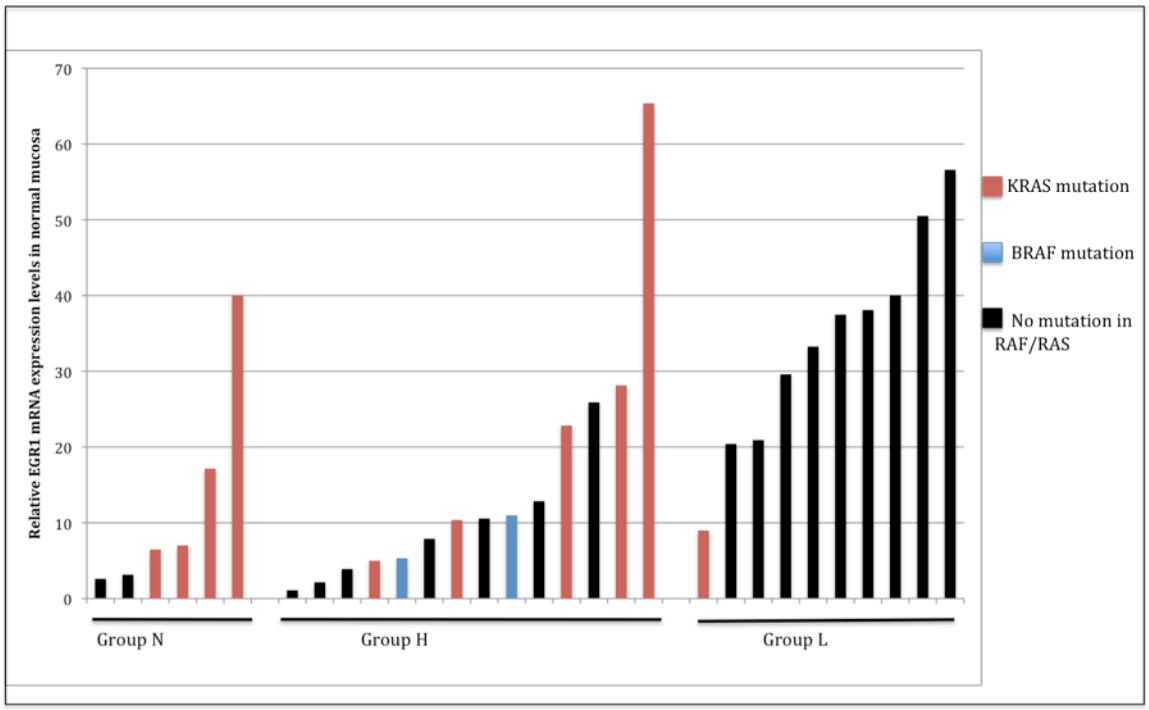
Table 3.2: *KRAS* and *BRAF* mutations in CRC patients

It was observed that 6/11 patients with either a *KRAS* or a *BRAF* mutation have higher expression of EGR1 in the tumour and the other 5/11 patients with *BRAF* and *KRAS* mutations are in the CRC patients that shows no significant difference between normal and tumour mucosa. Only one patient that shows decreased expression of EGR1 in the tumour harbouring *BRAF* or *KRAS* mutations. Therefore we looked to see if there was any association between the expression of EGR1 and the muation status of KRAS or BRAF.

It is clear from Figure 3.7, that patients with either a KRAS or BRAF mutation in the tumours have high levels of EGR1 in the tumour, and most belong to the group which shows low levels of EGR1 in the normal mucosa. The group of patients with EGR1 expression levels similar to those of the healthy control population have higher levels of EGR1 in the normal mucosa and signicantly lower levels in the tumour, with no KRAS or BRAF mutations. It could be therefore that presence of a KRAS/BRAF mutation is associated with low EGR1 levels in the normal mucosa with higher levels of EGR1 in the tumour.

In order to determine if the expression EGR1 correlates with either a BRAF or KRAS mutations, the expression graphs from Figure 3.4 were annotated as regards KRAS or BRAF mutation status (Figure 3.7). Graph A illustrates the EGR1 expression in the normal mucosa, with patients who have KRAS mutations in the matched tumour illustrated in red, with patients who have BRAF mutation in blue. Graph B illustrates the EGR1 expression in the tumour, with the corresponding KRAS and BRAF muations denoted in red and blue respectively. This figure further illustrates that KRAS and BRAF mutations are found in patients with higher levels of EGR1 in the tumour, with significantly lower lever of EGR1 in the normal mucosa.

A.


Relative EGR1 mRNA expression levels in normal mucosa
70
60
50
40
30
20
10
0
Group N
Group H
Group L
KRAS mutation
BRAF mutation
No mutation in RAF/RAS

B.

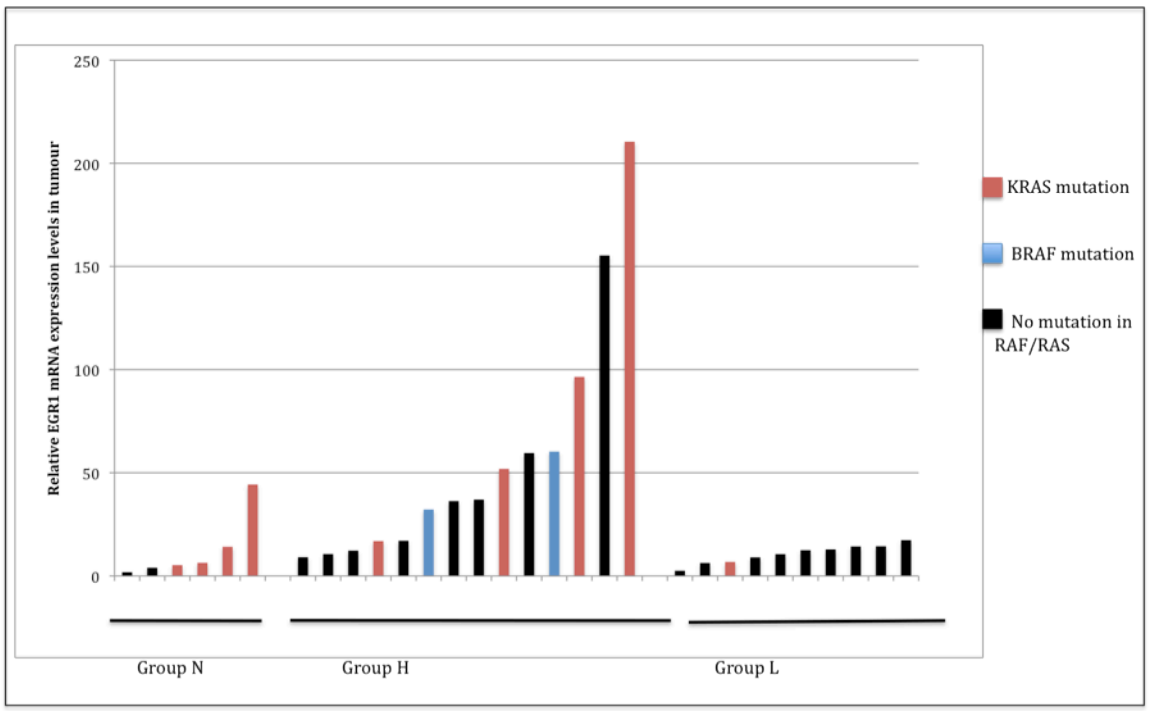

Relative EGR1 mRNA expression levels in tumour
250
200
150
100
50
0
Group N
Group H
Group L
KRAS mutation
BRAF mutation
No mutation in RAF/RAS

Figure 3.7: KRAS and BRAF mutation status with relative EGR1 mRNA expression levels in CRC patients.

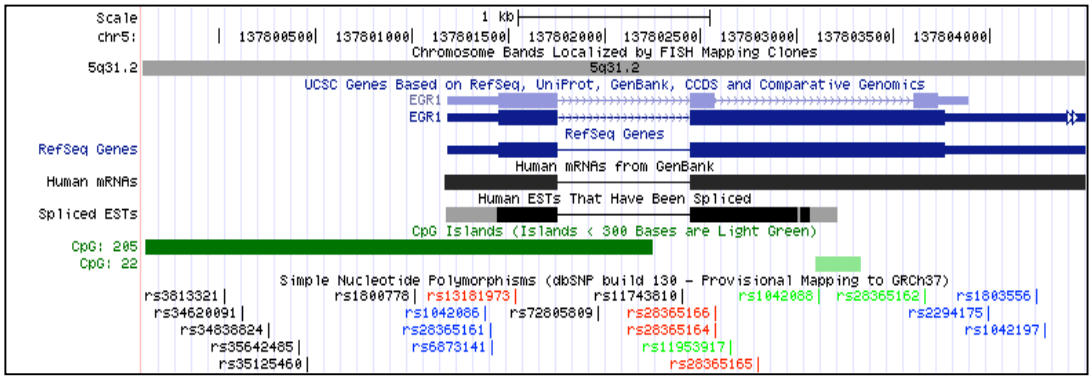
The data was divided a before based on difference in expression between normal mucosa and tumour, Group N have no difference in EGR1 expression levels, Group H have significantly higher levels in the tumour compared with normal and Group L have significantly lower levels of EGR1 in the tumour. **A.** Illustrates the EGR1 expression levels in the normal mucosa, patients with KRAS mutations in the matched tumour are denoted in red, patients with BRAF mutations in the tumour are denoted in blue. **B.** Illustrates the EGR1 expression levels in the tumour, with the presence of KRAF mutations in red and BRAF mutations in blue.

3.3.5 Methylation status of the *EGR1* promoter

Aberrant methylation of CpG islands in promoter regions frequently occurs in CRC and is associated with gene silencing, so the hypothesis that the differential expression of *EGR1* in the normal and tumour is associated with aberrant methylation of the *EGR1* promoter was tested as it may be that the group of patients that show much lower levels of *EGR1* in the tumour have aberrant methylation of the promoter. The methylation status of *EGR1* was determined using bisulfite sequencing of the *EGR1* promoter, which has a large CpG island, containing 205 CpG dinucleotides (Figure 3.7 A). Primers were designed using Methprimer (Li and Dahiya 2002), covering 80% of the CpG island and allowed sequencing of ~75% of the CpGs, and corresponds with the predicted CpG island using UCSC and the regions previously reporting *EGR1* promoter methylation (Seyfert, McMahon et al. 1990) . A normal mucosa sample was treated with a DNA methylase enzyme, SAM, as a positive control for the assay.

Bisulfite sequencing was conducted on the CRC cell lines and all of the 30 normal mucosa and tumour CRC patient samples. However no methylation of the *EGR1* promoter was found in any of these samples (Figure 3.7 B). Therefore the differential expression of *EGR1* that is observed is not associated with aberrant methylation of this region of the *EGR1* promoter.

A.



B.

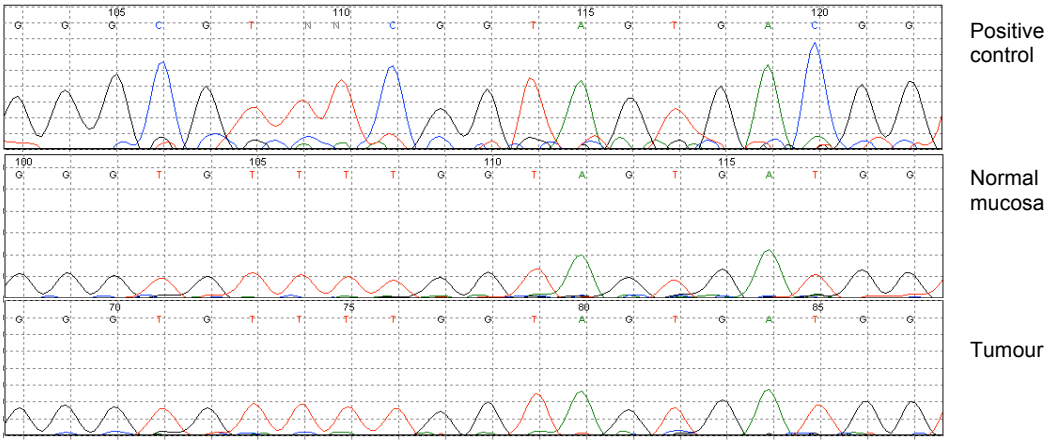


Figure 3.8: Bisulfite sequencing of matched normal and tumour samples

A. Primers were designed using MethPrimer (Li and Dahiya 2002) to cover the CpG island in the *EGR1* promoter as described in UCSC genome browser (highlighted in green). The primers covered 80% of the region and amplified 75% of the CpGs in the *EGR1* promoter. **B.** CRC patient normal and tumour DNA was isolated and bisulfite treated. The samples were amplified and sequenced to demonstrate any methylation in the *EGR1* promoter region. A sample was treated with S-(5'-Adenosyl)-L-homocysteine as a positive control for methylation as seen by the non-conversion of the C to T

3.3.6 Relative *EGR1* mRNA expression levels in IBD patients

The relative *EGR1* expression in IBD was investigated using qRT-PCR. Expression of *EGR1* in a healthy control population, as well as the two disease groups, ulcerative colitis patients and Crohn's disease patients was determined. The colonic biopsies were all taken from apparently un-inflamed tissue. An inflammatory condition was mimicked by culturing the biopsies with several different inflammatory mediators. The treatments used were LPS, MDP, PGN and TNF, all at a concentration of 1μg/ml and treated overnight.

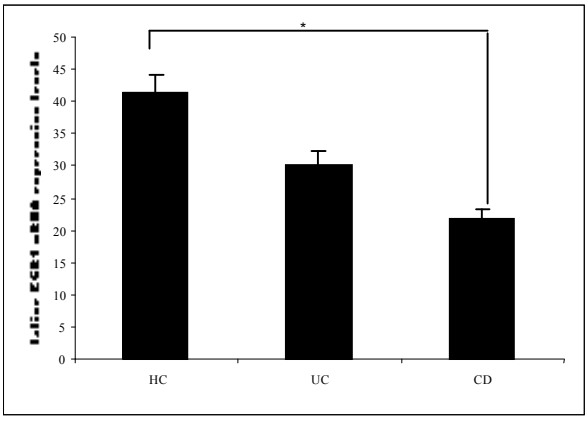
3.3.6.1 *EGR1* expression is significantly lower in IBD patients compared with healthy controls

The relative *EGR1* mRNA levels were examined in 17 healthy controls, 17 patients with Crohn's disease and 24 patients with ulcerative colitis to determine if there is any difference in the levels of *EGR1* in the three patient groups (Figure 3.8). By plotting the average of the *EGR1* levels for the three patient groups it demonstrates that there are considerably lower levels of *EGR1* expression in the both the ulcerative colitis and Crohn's disease patient groups. Using a mixed effect ANOVA model it was determined that these differences were statistically significant. Further analyses were preformed to determine which patient groups were significantly different from each other. There is a significant difference between the healthy controls and the Crohn's disease patients with a p value of 0.04437 (Figure 3.8 A). The decrease in *EGR1* expression is between the healthy controls and ulcerative colitis patients is not significant. There is also no significant difference between the Crohn's disease patients and the ulcerative colitis patients.

It should be noted that outliers were removed at a >3x the interquartile distance. There were no outliers present in the healthy controls but a single patient with Crohn's disease and two patients with ulcerative colitis were removed as outliers.

However, the values are still significant when the data is analysed with the outliers included. Therefore it is evident that there are significantly lower levels of *EGR1* being expressed in the colon of Crohn's disease patients compared with the healthy controls.

A.



B.

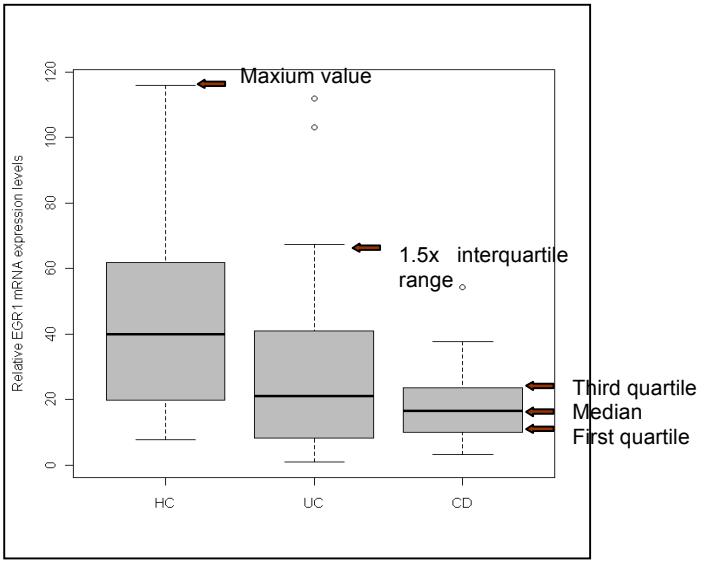


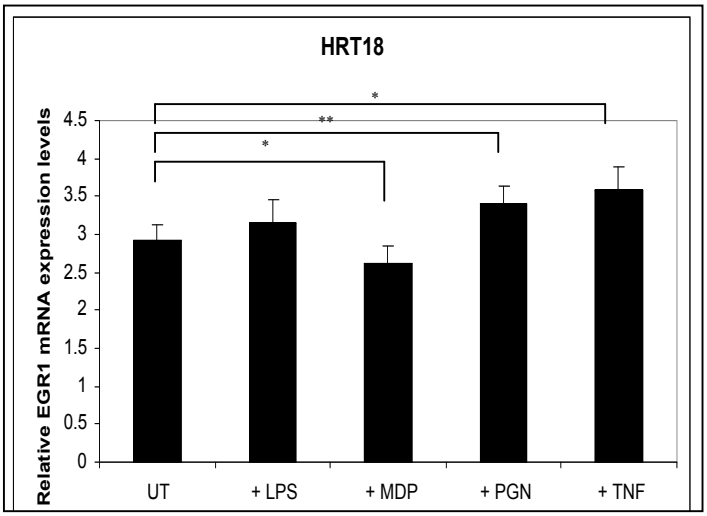
Figure 3.9: Relative *EGR1* mRNA expression levels in untreated IBD samples

RNA from biopsies of healthy controls (n=17), ulcerative colitis patients (n=24) and Crohn's disease patients (n=17) was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to *β-actin*. The experiment was performed in triplicate and repeated three times. **A**. The average value for each of the patient groups was plotted. The data was normalised using a rank transformation with the significance difference determined by a mixed effect ANOVA, significance was set at <0.05. **B**. A boxplot demonstrating the range of values obtained for the three patient groups, with the median and quartile distances illustrated. Outliers >3x the interquartile distance were removed from the analysis.

3.3.6.2 *EGR1* expression in CRC cell lines after inflammatory treatment

Having established that non-inflammed colonic tissue from the different patient groups expressed *EGR1* to different relative levels, we wanted to test the effect of inflammatory mediators of *EGR1* expression in those tissues. As *EGR1* is a stress response gene it was expected that the expression of *EGR1* would be induced by any treatment with inflammatory stimuli, and it is well documented that both LPS and TNF induce *EGR1* expression in several cell types (Cao, Guy et al. 1992; Coleman, Bartiss et al. 1992; Yao, Mackman et al. 1997; Kadl, Huber et al. 2002; Granet and Miossec 2004; Luyendyk, Schabbauer et al. 2008). The relative *EGR1* mRNA expression levels after treatment with LPS, MDP, PGN and TNF were initially looked at in the CRC cell lines HRT18 and SW480. The cells were treated with 1μg/ml of each inflammatory mediator for 24hrs, following which the RNA was extracted and the expression of *EGR1* determined by qRT-PCR. As previously found, the levels of *EGR1* in HRT18 are very low (Figure 3.9 A), and there appears to be induction of expression after PGN and TNF treatment. The expression of *EGR1* in SW480 cells is higher, and as expected there is an increase in *EGR1* expression after treatment will all of the inflammatory mediators, particularly MDP and TNF (Figure 3.9 B).

A.



B.

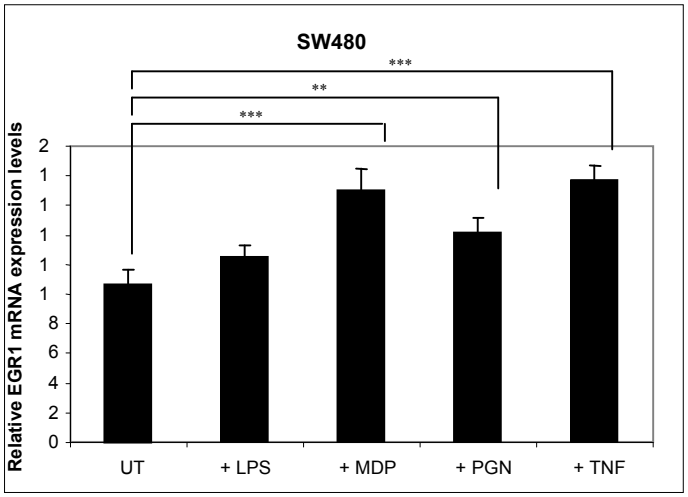


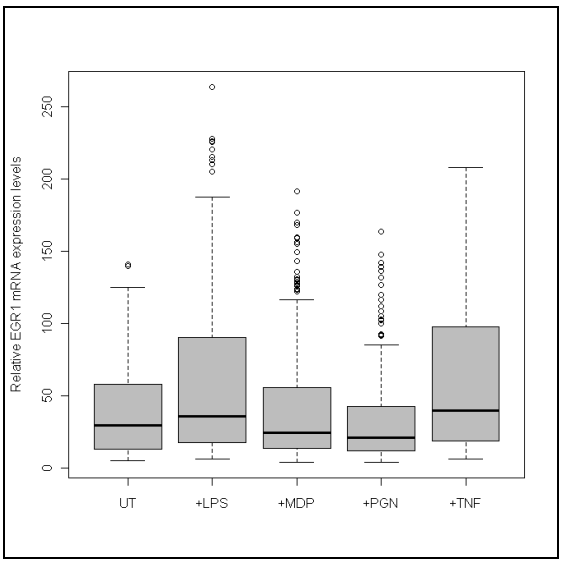
Figure 3.10: Relative *EGR1* mRNA expression levels in CRC cell lines after inflammatory treatment

HRT18 (**A**) and SW480 (**B**) CRC cell lines were treated with LPS, MDP, PGN and TNF (1 μg/ml) for 24hrs, the RNA was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to *β-actin*. The experiment was conducted in triplicate and repeated three times.

3.3.6.3 *EGR1* expression is significantly induced by inflammatory mediators in healthy controls

The relative *EGR1* mRNA levels after treatment by inflammatory mediators was first examined in the healthy control population to determine what effect the treatments would have on expression levels in healthy mucosa. In the healthy control patients *EGR1* expression appears to be induced by both LPS and TNF treatments, with very little difference evident after treatment with MDP and PGN, (Figure 3.10). The mixed effect ANOVA gave a p value of < 0.0001, and again as this was significant, multiple comparison tests were preformed. *EGR1* expression is significantly induced by both LPS and TNF in the healthy controls, with p values of $< 1x10^{-4}$ for both. MDP slightly induces *EGR1* expression although not to a significantly different level. PGN is the only treatment that does not appear to induce any *EGR1* expression (Figure 3.10 B). This suggests that *EGR1* is induced in normal mucosa as part of an inflammatory response to both LPS and the TNF cytokine.

A.



B.

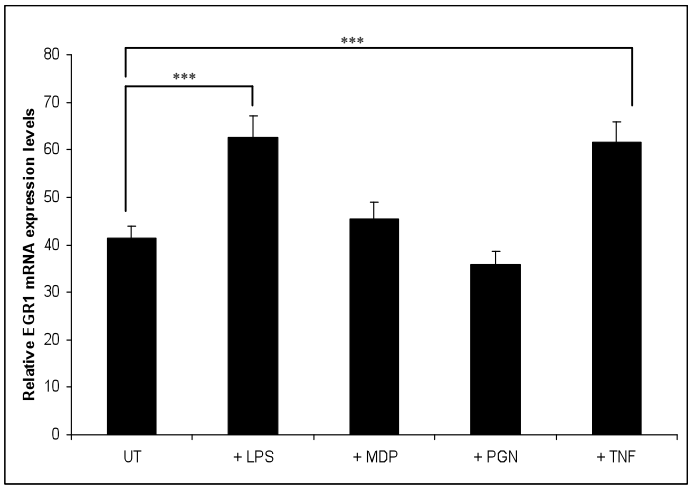


Figure 3.11: Relative *EGR1* mRNA expression levels in healthy controls

Biopsies of healthy controls (n=17) was isolated and the relative mRNA expression levels were treated with LPS, MDP, PGN and TNF (1μg/ml) for 24hrs, the RNA was extracted and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to *β-actin*. The experiment was performed in triplicate and repeated three times. **A.** A boxplot demonstrating the range of values obtained for the five treatments, with the median and quartile distances illustrated. **B.** The average value for each of the patient groups was plotted. The data was normalised using a rank transformation with the significance difference determined by a mixed effect ANOVA, significance was set at <0.05. Outliers >3x the interquartile distance were removed from the analysis.

In order to look more closely at the individual patients rather than the overall population, each patient was analysed using Wilcoxon rank sum test, comparing each treatment against the untreated (Table 3.3). Figure 3.11 A illustrates the effect of treatment with each inflammatory mediator, showing the percentage of how the healthy controls were affected. Using the Wilcoxon test, each patient was grouped based on the effect, i.e. if the treatment caused *EGR1* to be increased, increased significantly, decreased, decreased significantly or showed little or no effect at all. This correlates quite well with the population analysis, as it is clear that the majority of patients show a significant increase after both LPS and TNF treatment, whereas patients show a decrease as well as an increase after MDP treatment.

Figure 3.11 B illustrates the fold increase or decrease of each patient after each treatment. The untreated value is set 1, with the treated values adjusted accordingly and plotted on a logarithmic scale. It is clear from this graph that for the majority of patients who show increased levels of *EGR1* after treatment do so after treatment with all the inflammatory mediators, as in most cases if there is an increase in *EGR1* after LPS, there is an increase after MDP, PGN and TNF treatment also. The same is true for the patients that show a decrease in *EGR1* after treatment. There are few exceptions, and it would be interesting to determine why some patients respond in the opposite manner to the majority, i.e. those that show a decrease in *EGR1* expression after one treatment.

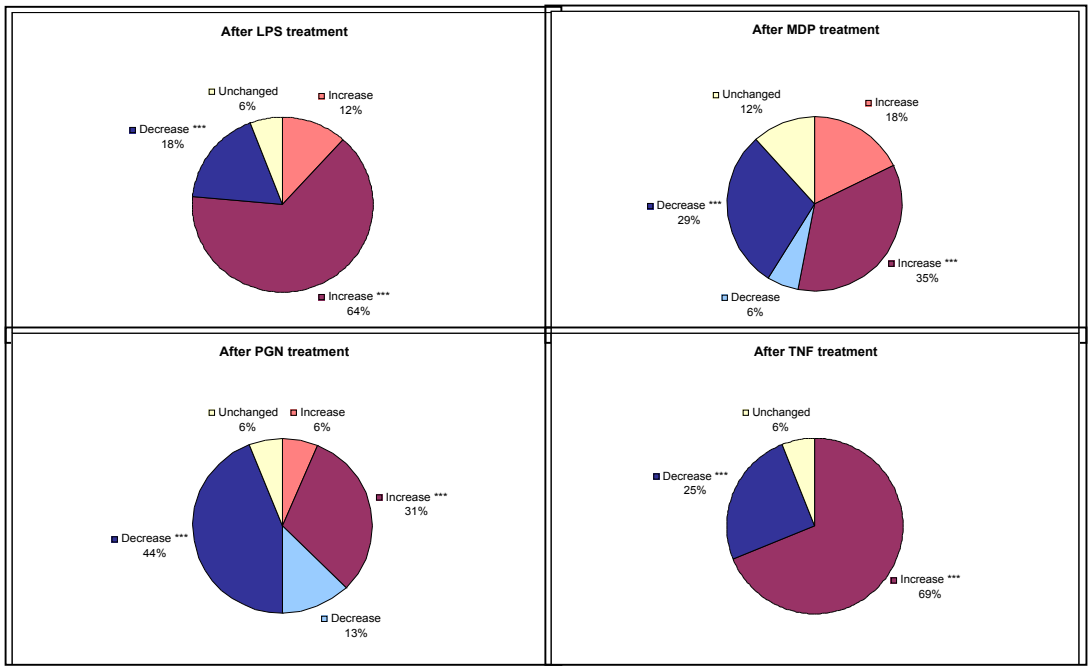
Healthy Controls

	+ LPS		+ MDP		+ PGN		+ TNF	
	Increase	Decrease	Increase	Decrease	Increase	Decrease	Increase	Decrease
5JN	0.06396		0.0004883 ***			0.03418 *		
9CD	0.003906 **		0.09766			0.003906*	No change - 0.6523	
13CS	0.003906 **			0.003906 **	0.003906 **		0.003906 **	
17GF		0.007812 **		0.003906 **		0.003906 **		0.003906 **
24PK	0.1641			0.003906 **		0.003906 **		0.003906 **
26MW		0.007812 **		0.003906 **		0.003906 **		0.003906 **
34CM	0.01172 *			0.003906 **		0.09766	0.007812 **	
35WM	0.003906 **		0.003906 **			0.25	0.003906 **	
38GW	0.003906 **		0.02734 *		0.003906 **		0.003906 **	
39TT	0.003906 **		0.003906 **		0.003906 **		0.003906 **	
42FD	0.02734 *		0.09766			0.003906 **	0.003906 **	
44JH	No change - 0.4238		No change - 0.129		No change - 0.1294		0.0004883 ***	
48EC	0.0004883 ***		No change - 0.4238		0.0004883 ***		0.0004883 ***	
50JB-S	0.01953 *		0.25		0.2031		0.007812 **	
51AW	0.003906 **		0.003906 **		0.003906 **		0.003906 **	
55HC	0.003906 **		0.003906 **			0.02734 *	0.003906 **	
57HC		0.003906 **		0.003906 **				0.003906 **
62HC	0.003906 **			0.09766	0.003906 **		0.003906 **	

* =p<0.05
** =p<0.01
*** =p<0.001

Table 3.3: Wilcoxon rank sum analysis of healthy control tissue after treatment with inflammatory mediators

A.



B.

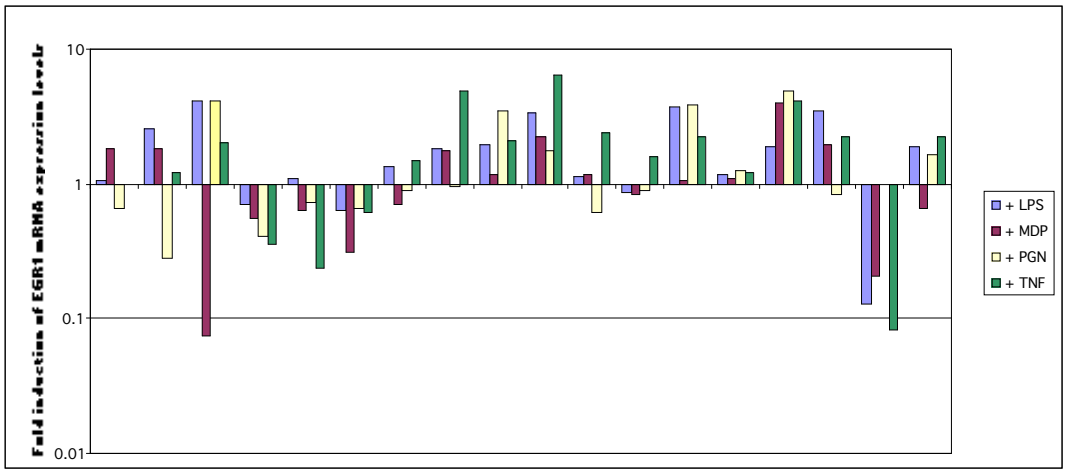


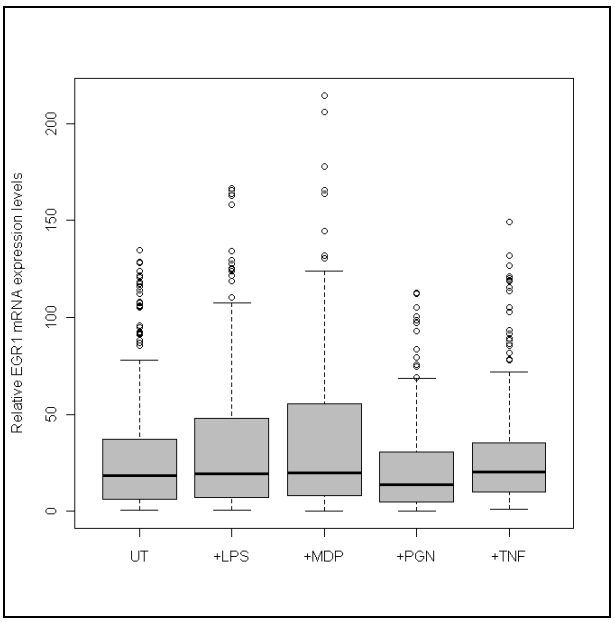
Figure 3.12: Fold induction of *EGR1* after treatment of healthy control tissue

An illustration of the effect of treatment with inflammatory mediators on *EGR1* expression levels in healthy controls. **A**. Demonstrates the proportion of healthy controls (n=17) showing an increase or decrease in *EGR1* after treatment. Each patient analysis was done using the Wilcoxon rank sum test (paired with untreated), with significance set at <0.05. **B**. The untreated biospy values were set to 1, with each group representing a patient after the four treatments. The plot illustrates the fold increase or decrease in *EGR1* expression after treatment with inflammatory mediators in the healthy controls.

3.3.6.4 *EGR1* expression is induced by inflammatory mediators in ulcerative colitis patients

In the ulcerative colitis patients the effect of the inflammatory mediators seems similar to that of the healthy control, except after PGN treatment, which actually appears to decrease the level of *EGR1* expression in these patients (Figure 3.12). The mixed effect ANOVA test gave a p value of < 0.0001, and once analysed further it showed that although LPS and MDP treatment appear to induce *EGR1* expression it is not to a significant level (Figure 3.12 B). There is again a significant increase in *EGR1* expression after treatment with TNF ($p = 0.00171$). However the PGN treatment causes the *EGR1* level in these patients to significantly decrease ($p = <0.0001$).

A.



B.

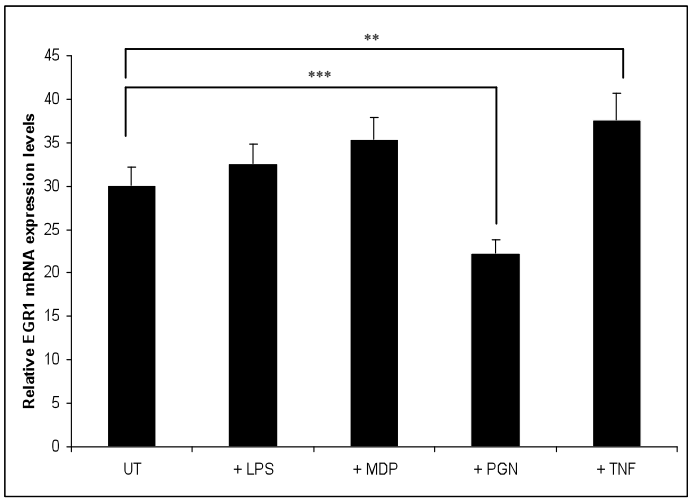


Figure 3.13: Relative *EGR1* mRNA expression levels in ulcerative colitis patients

RNA from biopsies of ulcerative colitis patients (n=24) was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to *β-actin*. The experiment was performed in triplicate and repeated three times. **A**. A boxplot demonstrating the range of values obtained for the five treatments, with the median and quartile distances illustrated. **B**. The average value for each of the patient groups was plotted. The data was normalised using a rank transformation with the significance difference determined by a mixed effect ANOVA, significance was set at <0.05. Outliers >3x the interquartile distance were removed from the analysis.

Investigating the patients on an individual basis indicates that the majority of the patients show a significant increase after TNF treatment, correlating with the population analysis, whereas after LPS and MDP treatment approximately half the patients show a decrease in *EGR1* expression after treatment (Table 3.4; Figure 3.13 A). Unlike the healthy controls only ~45% show an increase in *EGR1* after LPS treatment. After PGN treatment, the majority of patients show a decrease in *EGR1* expression, which is again significant in the population analysis of the UC patients. Hence when the fold induction is plotted for the UC patients the results demonstrate a different pattern to the HC graph (Figure 3.13 B). A few patients do show similarity to the HC samples in that if *EGR1* increases after one treatment, it increases after all, however this is not the case for most of the patients and indicates that the patients are not responding to all the treatments and inflammation in the same way.

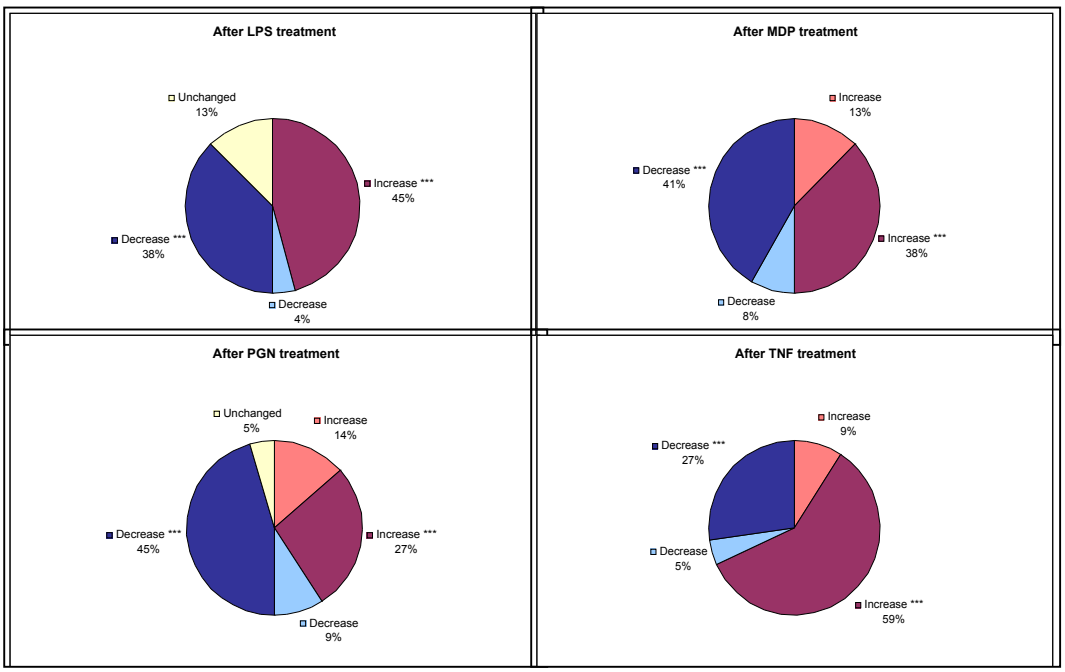
Ulcerative Colitis

	+ LPS		+ MDP		+ PGN		+ TNF	
	Increase	Decrease	Increase	Decrease	Increase	Decrease	Increase	Decrease
2HF	0.0009766 ***		0.0009766 ***			0.05371		0.0009766 ***
4PM	0.003906 **		0.007812 **			0.003906 **	0.003906 **	
6CP		0.0009766 ***		0.0009766 ***	No change -0.5771			
8BS		0.9102		0.02734 *		0.05469		0.1641
12DK	0.007812 **			0.003906 **	0.003906 **		0.003906 **	
13CS	0.003906 **			0.003906 **	0.003906 **		0.003906 **	
18AB		0.003906 **	0.003906 **		0.003906 **		0.003906 **	
19SU		0.003906 **	0.09766			0.003906 **	0.003906 **	
21RR	0.003906 **			0.003906 **		0.003906 **	0.003906 **	
22JC	No change - 0.9102		0.05469				0.003906 **	
23JG	No change - 0.1614		0.003906 **			0.007812 **	0.003906 **	
25HW		0.0002441 ***		0.1531		0.04944 *		0.04944 *
29GM		0.003906 **		0.003906 **		0.006104 **		0.003906 **
33SK	0.01343 *		0.008057 **		0.00524 **		1	
36GE		0.003906 **		0.007812 **		0.01172 *		0.003906 **
40AG	0.003906 **		0.003906 **			0.1289		0.1641
41WB	0.003906 **			0.003906 **	0.007812 **		0.25	
45RB	0.004883 **		0.002441 **			0.004883 **	0.004883 **	
47WD	0.01172 *		0.09766		0.1641		0.003906 **	
52DB		0.007812 **		0.003906 **	0.003906 **		0.003906 **	
54SK	0.003906 **		0.003906 **					
63UC	0.003906 **		0.003906 **		0.05469		0.003906 **	
64UC	No change	- 0.5703		0.003906 **	0.1289			0.003906 **
65UC		0.003906 **		0.25		0.003906 **	0.003906 **	
66UC		0.003906 **		0.003906 **		0.003906 **		0.003906 **

* =p<0.05
** =p<0.01
*** =p<0.001

Table 3.4: Wilcoxon rank sum analysis of Ulcerative colitis patient tissue after treatment with inflammatory mediators

A.



B.

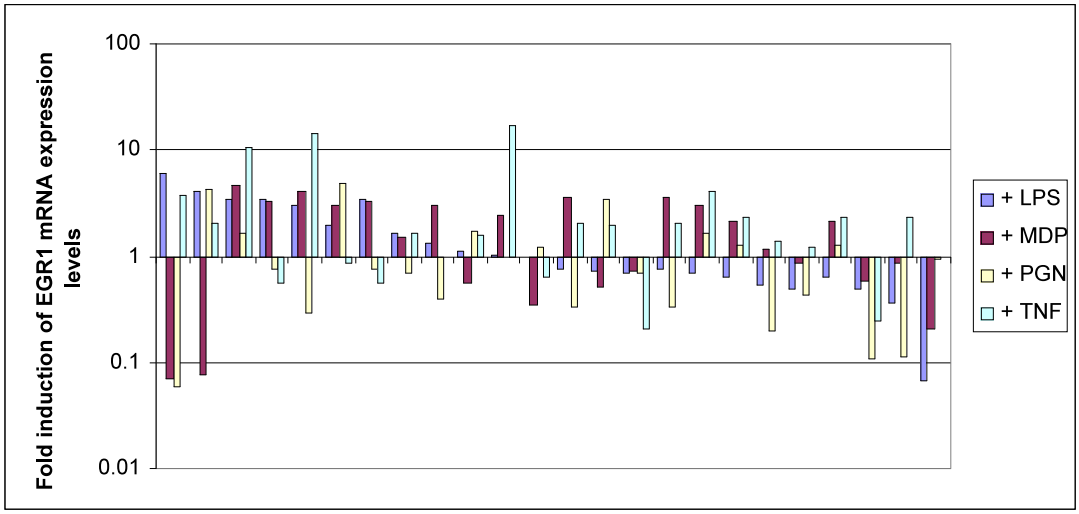


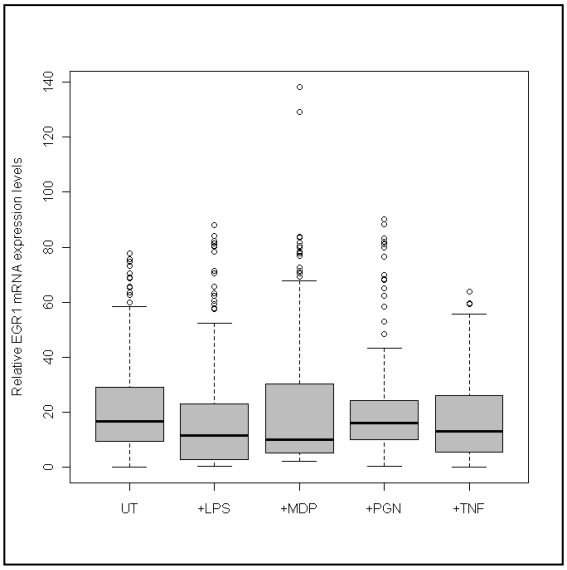
Figure 3.14: Fold induction of *EGR1* after treatment of ulcerative colitis tissue

An illustration of the effect of treatment with inflammatory mediators on *EGR1* expression levels in ulcerative colitis patients. **A.** Demonstrates the proportion of ulcerative colitis patients (n=24) showing an increase or decrease in *EGR1* after treatment. Each patient analysis was done using the Wilcoxon rank sum test (paired with untreated), with significance set at <0.05. **B.** The untreated was set a value of 1 to illustrate the fold increase or decrease in *EGR1* expression after treatment with inflammatory mediators in the ulcerative colitis patients.

3.3.6.5 *EGR1* expression is reduced after treatment with inflammatory mediators in Crohn's disease patients

The effect of the inflammatory mediators on the patients with Crohn's disease showed a different pattern to the HC and UC samples. In the CD patient the expression of *EGR1* is decreased after treatment with LPS and TNF compared with the untreated samples (Figure 3.14). Similar to the healthy controls, there appears to be little difference after treatment with MDP and PGN, but the other two treatments decrease the expression of *EGR1*. The mixed model ANOVA test gave a p value of < $1x10^{-4}$, and the multiple comparison tests showed that *EGR1* is significantly reduced after LPS ($p = 0.00619$). Interestingly although TNF treatment appears to reduce the expression of *EGR1* in the population analysis, when looked at on an individual basis, this is not reflected in the ANOVA analysis. However 66% of patients show a decrease in *EGR1*expression after TNF treatment, with 46% of these patients showing a significant decrease (Table 3.5; Figure 3.15 A). Similarly 65% of patients show a significant decrease after treatment with MDP, but there is no difference in the average levels of *EGR1* after MDP treatment in the population analysis. Similar to the HC population, the majority of patients respond similarly to all treatments.

A.



B.

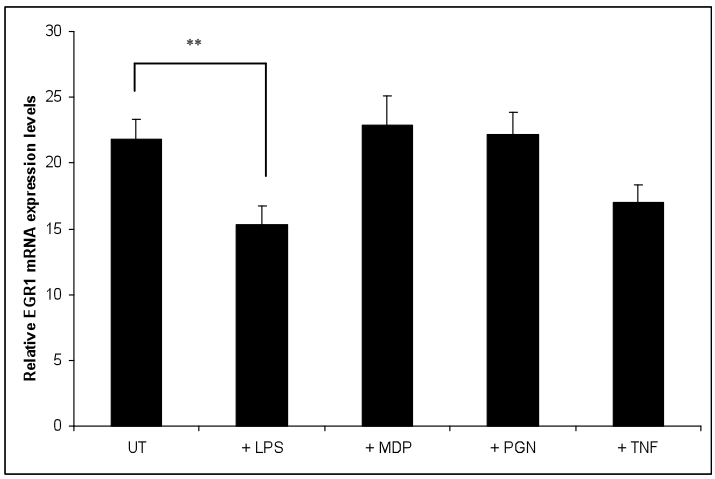


Figure 3.15: Relative *EGR1* mRNA expression levels in Crohn's disease patients

RNA from biopsies of Crohn's disease patients (n=17) was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *EGR1* expression levels relative to *β-actin*. The experiment was performed in triplicate and repeated three times. **A.** A boxplot demonstrating the range of values obtained for the five treatments, with the median and quartile distances illustrated. **B.** The average value for each of the patient groups was plotted. The data was normalised using a rank transformation with the significance difference determined by a mixed effect ANOVA, significance was set at <0.05. Outliers >3x the interquartile distance were removed from the analysis.

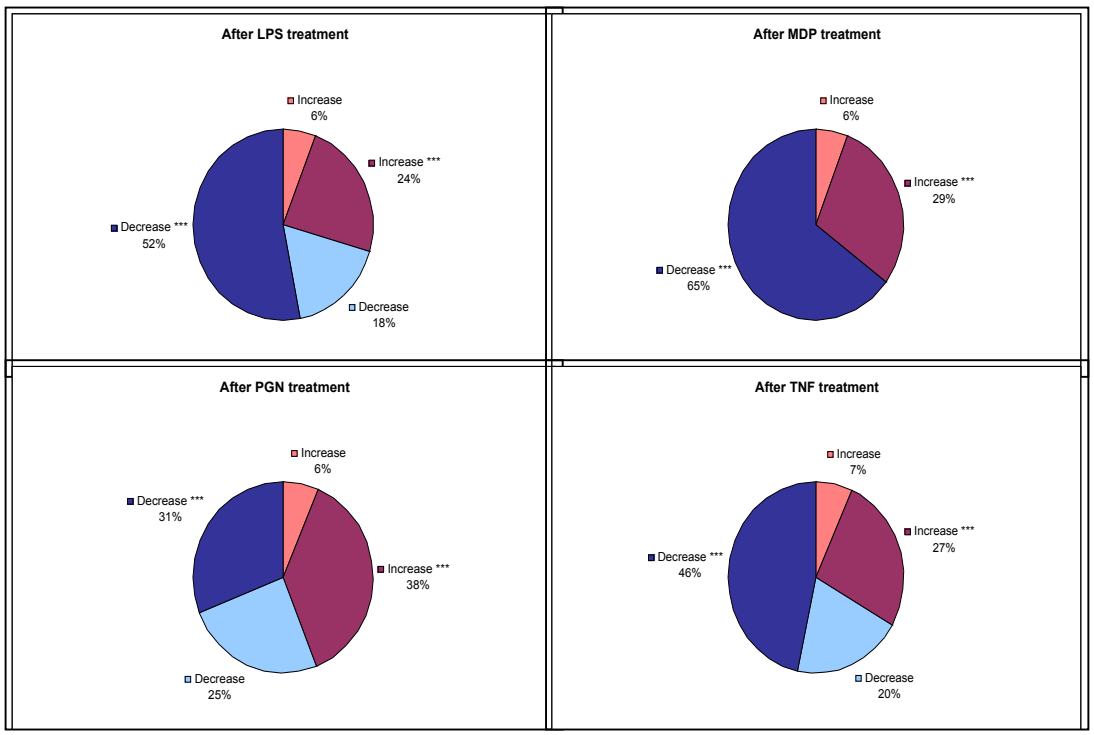
Crohn's Disease

	+ LPS		+ MDP		+ PGN		+ TNF	
	Increase	Decrease	Increase	Decrease	Increase	Decrease	Increase	Decrease
3NB	0.007813 **		0.007813 **		0.007813 **			
7LH		0.0009766 ***		0.004883 **		0.0009766 ***		0.0009766 ***
11SF		0.003906 **		0.003906 **		0.25	0.25	
14TS		0.003906 **		0.003906 **		0.003906 **		0.003906 **
15MR		0.07422	0.01953 *			1		0.4961
16JS		0.3594		0.003906 **		0.05469	0.01953 *	
20KS	0.003906 **			0.003906 **		0.003906 **		
27KH		0.004883 **		0.004883 **	0.01611 *			0.004883 **
31AM	0.003906 **		0.02734 *		0.05469		0.007812 **	
32MC	0.3594		0.4258		0.007812 **		0.003906 **	
46RP		0.0009766 ***	0.0009766 ***		0.04199 *			0.0009766 ***
53LL		0.003906 **		0.003906 **		0.07422		0.3008
56CD		0.003906 **		0.01953 *	0.003906 **			0.003906 **
58CD		0.003906 **		0.003906 **		0.003906 **		0.003906 **
60CD	0.02734 *		0.003906 **		0.003906 **		0.003906 **	
61CD		0.003906 **		0.003906 **		0.01172 *		0.003906 **

* =p<0.05
** =p<0.01
*** =p<0.001

Table 3.5: Wilcoxon rank sum analysis of Crohn's disease patient tissue after treatment with inflammatory mediators

A.



B.

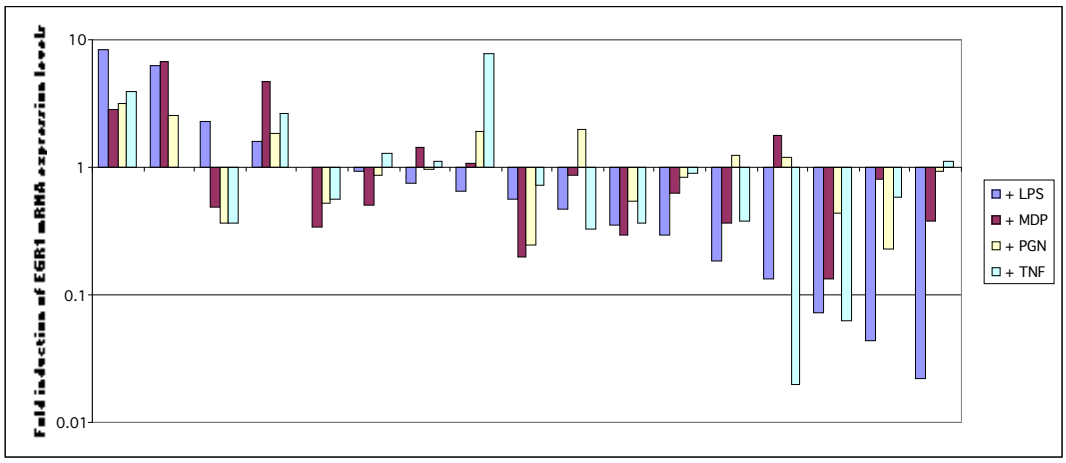


Figure 3.16: Fold induction of *EGR1* after treatment of Crohn's disease tissue

An illustration of the effect of treatment with inflammatory mediators on *EGR1* expression levels in Crohn's disesase. **A**. Demonstrates the proportion of Crohn's disease (n=17) showing an increase or decrease in *EGR1* after treatment. Each patient analysis was done using the Wilcoxon rank sum test (paired with untreated), with significance set at <0.05. **B**. The untreated biopsy samples were set to a value of 1 to illustrate the fold increase or decrease in *EGR1* expression after treatment with inflammatory mediators in the Crohn's disease.

3.3.7 Genotype analysis of IBD patients for *EGR1* variants

To address whether *EGR1* variation was affecting gene expression the IBD patient samples were genotyped for the three *EGR1* variants in order to determine if there is any correlation with *EGR1* variants and the lower levels of *EGR1*in both Crohn's disease and ulcerative colitis (Table 3.6). Figure 3.16 A shows that there is no difference in the frequency of the alleles in the three *EGR1* SNPs in the patient groups, nor is there any difference in gene expression with the presence of the different alleles in any of the three SNPs (Figure 3.16 B). Interestingly, however, in the healthy controls, there is only 1 patient with the homozygous wild-type genotype. The healthy controls have the highest EGR1 expression, with expression levels similar to that of the group L, from the CRC patients, where there was no wild-type genotypes present (Figure 3.4, 3.5).

Healthy Controls

	rs3813221	rs11743810	rs11748288
5JN	CC	CC	AA
26MW	CT	CT	AG
34CM	CT	CT	AG
35WM	CT	CT	AG
44JH	CT	CT	AG
50JB-S	CT	CT	AG
51AW	CT	CT	AG
55HC	CT	CT	AG
62HC	CT	CT	AG
9CD	TT	TT	GG
17GF	TT	TT	GG
24PK	TT	TT	GG
42FD	TT	TT	GG
67HC	TT	TT	GG
13CS	CT	TT	GG
38GW	CT	TT	GG
39TT	CC	CT	AG
57HC	CT	TT	GG

Ulcerative Colitis

	rs3813221	rs11743810	rs11748288
6CP	CC	CC	AA
18AB	CC	CC	AA
36GE	CC	CC	AA
49UC	CC	CC	AA
8BS	CT	CT	AG
2HF	CT	CT	AG
4PM	CT	CT	AG
19SU	CT	CT	AG
29GM	CT	CT	AG
33SK	CT	CT	AG
41WB	CT	CT	AG
47WD	CT	CT	AG
52DB	CT	CT	AG
54SK	CT	CT	AG
66UC	CT	CT	AG
22JC	TT	TT	GG
23JG	TT	TT	GG
63UC	TT	TT	GG
65UC	TT	TT	GG
12DK	CC	TT	GG
21RR	CT	TT	GG
25HW	CC	CT	AG
40AG	TT	TT	AG
45RB	CT	TT	GG
64UC	CC	CT	AG

Crohn's Disease

	rs3813221	rs11743810	rs11748288
28CD	CC	CC	AA
31AM	CC	CC	AA
56CD	CC	CC	AA
7LH	CT	CT	AG
32MC	CT	CT	AG
58CD	CT	CT	AG
3NB	TT	TT	GG
20KS	TT	TT	GG
61CD	TT	TT	GG
11SF	CC	CT	AG
14TS	CT	TT	GG
15MR	CC	CT	AG
16JS	TT	CT	GG
27KH	CT	TT	GG
46RP	CT	TT	GG
53LL	CC	TT	GG

Table 3.6: Genotype of IBD and healthy control patients

A. B.

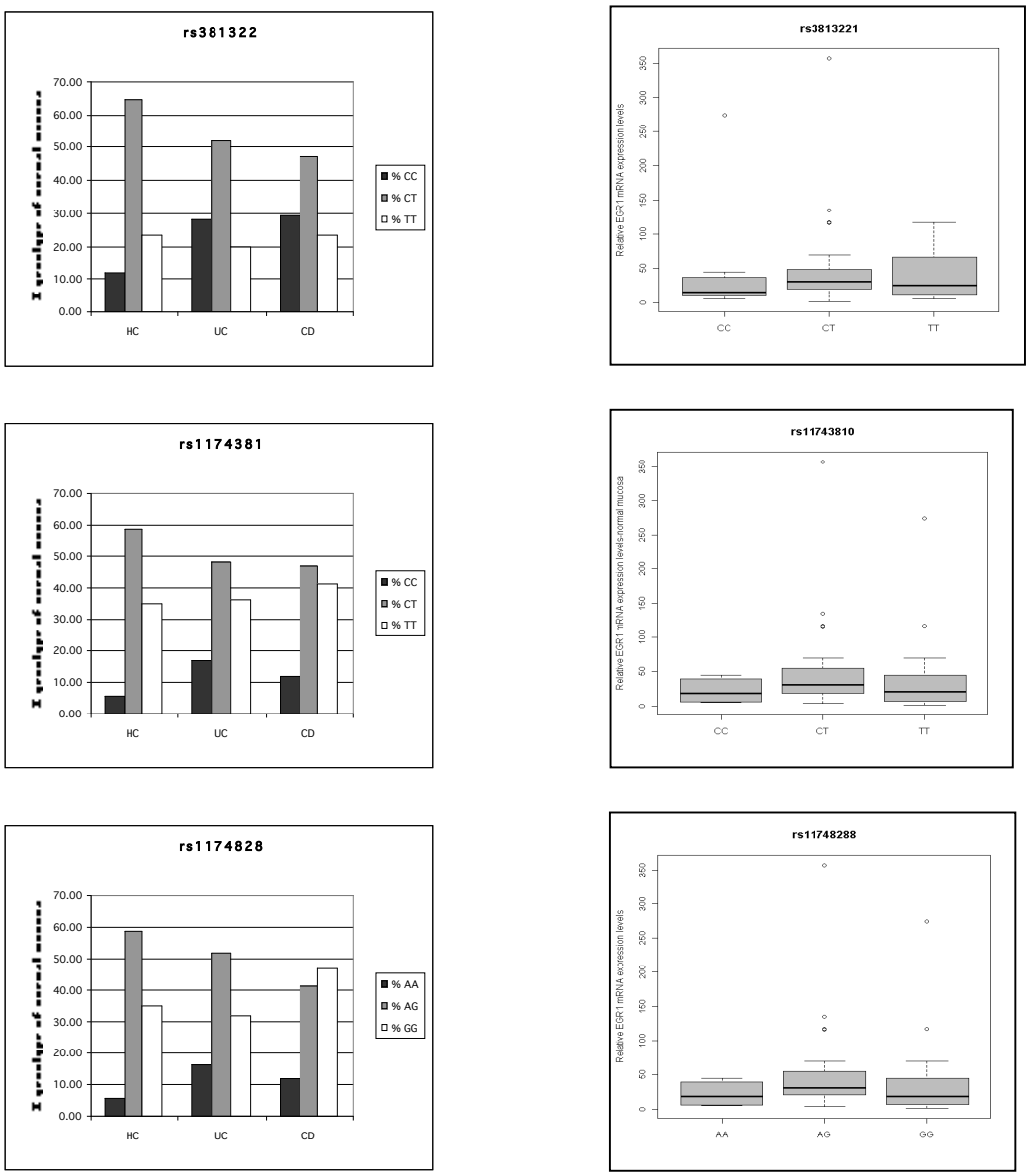


Figure 3.17: Genotyping of IBD patients and healthy controls

DNA from biopsies of healthy controls (n=17), ulcerative colitis patients (n=24) and Crohn's disease patients (n=17) was isolated, amplified by PCR and sequenced to determine the genotype of three *EGR1* variants. **A.** Illustrates what percentage of alleles is present for each of the three SNPs in the three patient groups. **B.** The relative *EGR1* expression levels are plotted against the alleles of the three SNPs to determine if there is a correlation with presence of the *EGR1* variants and gene expression.

3.3.8 Methylation status of the *EGR1* promoter in IBD

The *EGR1* promoter was investigated to determine if there was any methylation of the CpG's that may correlate with the differential gene expression in the patient groups. The DNA from all of the healthy controls, ulcerative colitis patient and Crohn's disease patients was bisulfite treated and sequenced at the *EGR1* promoter as in section 3.3.5. There was no CpG methylation evident at the *EGR1* promoter in any of the samples (Figure 3.17).

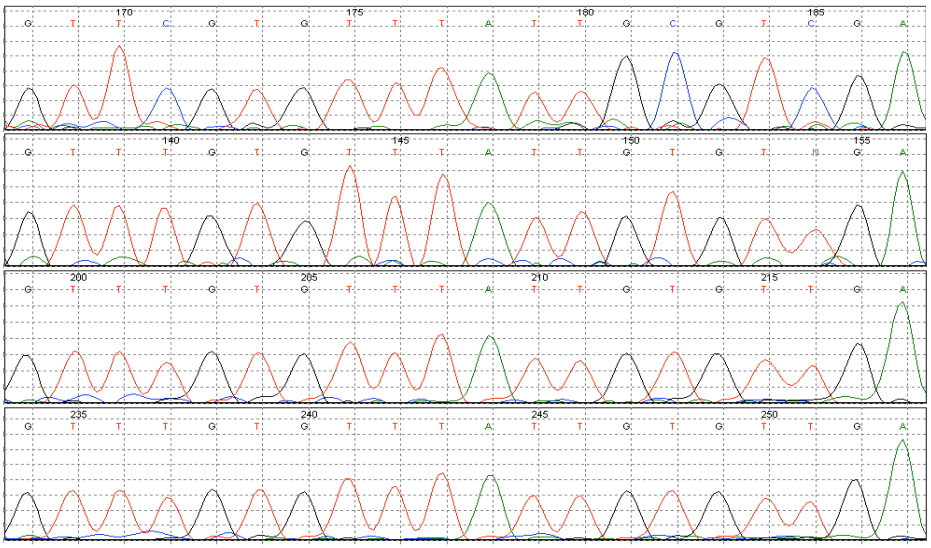


Figure 3.18: Bisulfite sequencing of healthy controls and IBD patients

DNA was isolated from all the healthy controls, ulcerative colitis patients and Crohn's disease patients and bisulfite treated. The samples were amplified and sequenced to demonstrate any methylation in the *EGR1* promoter regions. A sample was treated with S-(5'-Adenosyl)-L-homocysteine as a positive control.

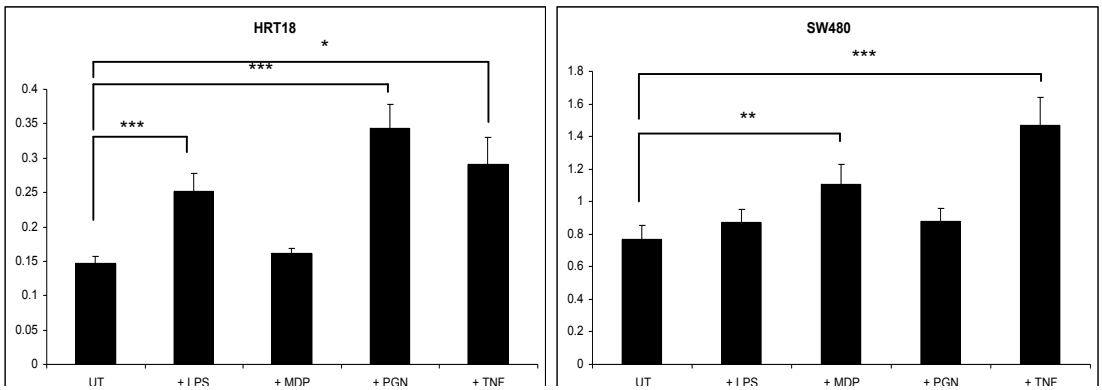
3.3.9 *NAB2* mRNA expression

One of the main mechanisms by which the expression of *EGR1* is regulated is by means of two co-repressors. Both NAB1 and NAB2 can block the biological activity of EGR1. NAB1 is a constituently expressed protein but NAB2 has been reported to be activated by the same stimuli as *EGR1* and can be activated by EGR1 itself (Kumbrink, Gerlinger et al. 2005). It therefore acts in a negative feedback loop with EGR1. One hypothesis is that in the Crohn's disease patients, the inflammatory condition of the disease is over-stimulating the expression of the NAB2 inhibitor, leading to lower levels of *EGR1* in the patients and a dysregulation of the feedback loop. This hypothesis was tested by examining expression of *NAB2* in the IBD patients and healthy controls.

3.3.9.1 The relative *NAB2* expression levels in CRC cell lines

The NAB2 mRNA expression levels were first examined in CRC cell lines HRT18 and SW480 after treatment with the inflammatory mediators in order to ensure that NAB2 expression could be detected. Figure 3.18 A. illustrates that the levels of endogenous *NAB2* in both cell lines is very low and although there does appear to be significant induction of expression after treatment in the cell lines this does not cause the *NAB2* levels to greatly increase overall when compared with the *EGR1* levels (Figure 3.18 B)

A.



B.

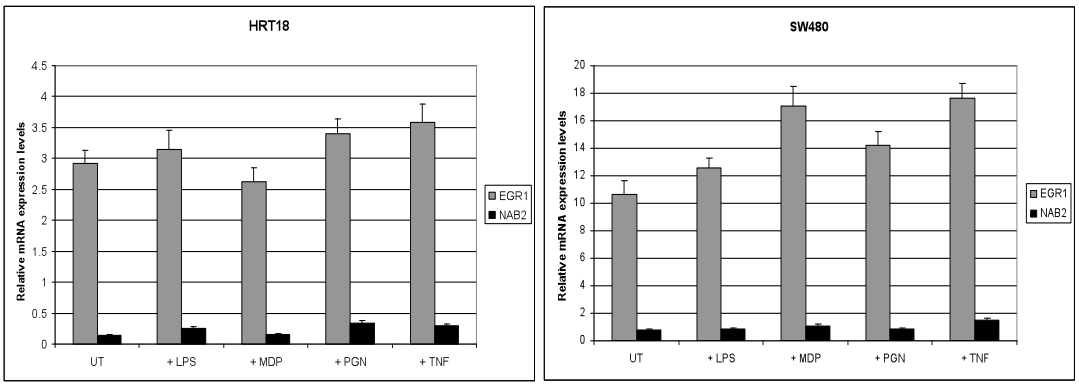


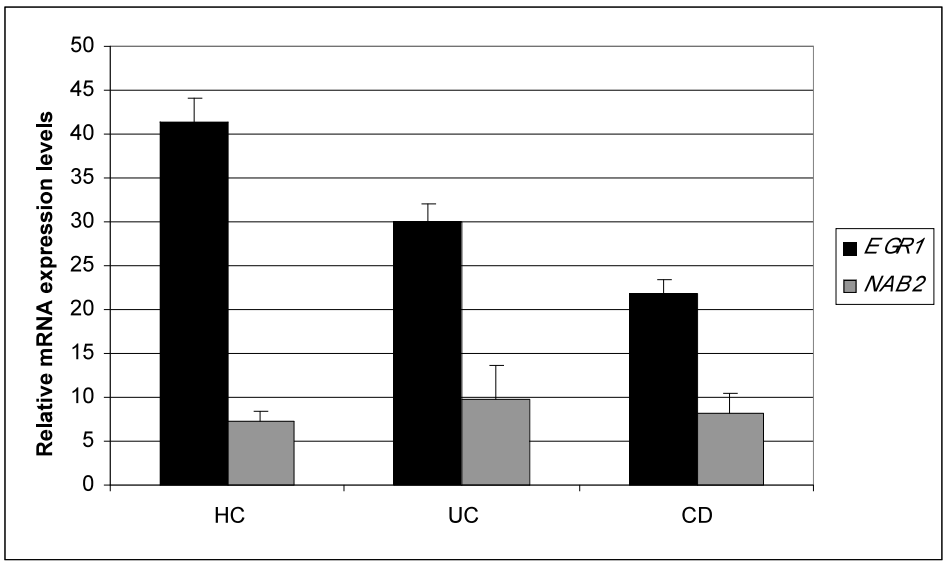
Figure 3.19: Relative *NAB2* mRNA expression in CRC cell lines after inflammatory treatment

HRT18 and SW480 CRC cell lines were treated with LPS, MDP, PGN and TNF (1 μg/ml) for 24hrs, the RNA was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The experiment was conducted in triplicate and repeated three times. The values show *NAB2* expression levels relative to *β-actin*. **A.** The relative *NAB2* mRNA expression levels in HRT18 and SW480 cell, untreated and after treatment with inflammatory mediators. **B.** *EGR1* and *NAB2* expression levels in HRT18 and SW480 cells.

3.3.9.2 Relative *NAB2* expression levels in IBD

The relative *NAB2* mRNA expression levels were examined in healthy controls (n=19), ulcerative colitis patients (n=19) and patients with Crohn's disease (n=14) representing 89.89%, 76% and 77.78% respectively of the total patients in each group. Figure 19.A. shows that there are relatively low levels of *NAB2* present, and much lower levels of *NAB2* compared with *EGR1* in all three of the patient groups. Figure 19.B illustrates the level of *NAB2* compared to *EGR1* (where *EGR1* value=1), plotted on a logarithmic scale. In the healthy controls none of the patients have higher levels of *NAB2* than *EGR1*. Three patients have higher levels of *NAB2* in ulcerative colitis (15.79%) and two Crohn's disease patients have higher levels of *NAB2* compared with *EGR1* (14.29%). There appears to be no differential expression of *NAB2* in ulcerative colitis or Crohn's disease although the patient numbers are small and the low expression of *NAB2* may be an issue.

A.



B.

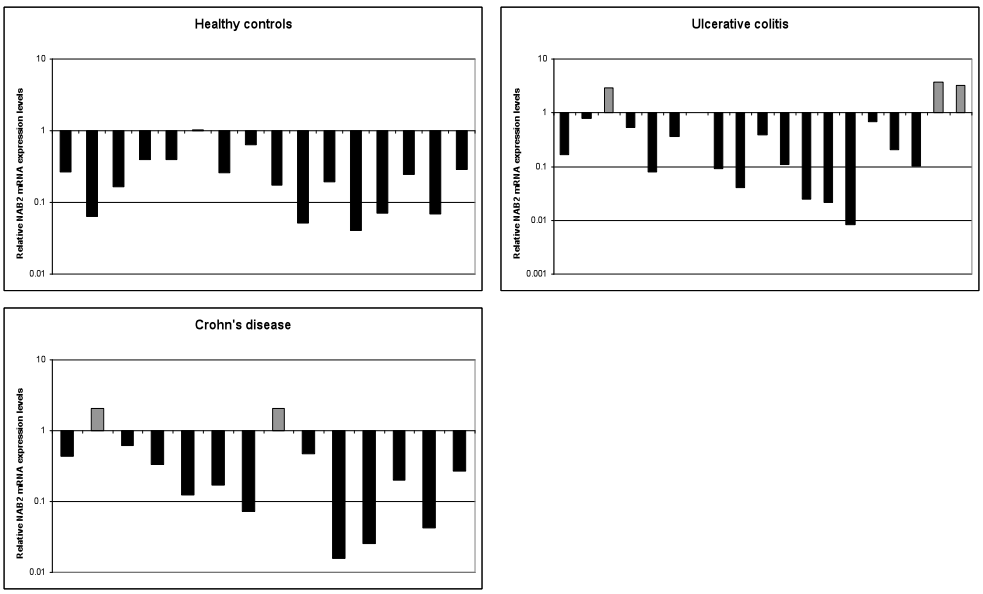


Figure 3.20: Relative *NAB2* mRNA expression levels in IBD and healthy controls

RNA from biopsies of healthy controls (n=16), ulcerative colitis patients (n=19) and Crohn's disease patients (n=14) was isolated and the relative mRNA expression levels were determined by quantitative real-time PCR. The values show *NAB2* and *EGR1* expression levels relative to *β-actin*. The experiment was performed in triplicate and repeated twice. **A.** The average value of *NAB2* expression for each of the patient groups was plotted, along with the average value of *EGR1* expression. **B.** The *EGR1* values were set to 1 to illustrate the fold increase or decrease in endogenous *NAB2* expression compared with the endogenous *EGR1* levels for each patient.

3.4 Discussion

It is clear that there is differential expression of *EGR1* in diseased colon. We looked at the expression of *EGR1* in matched normal mucosa and tumour and found that there is significantly different expression in 23/30 CRC patients. Given that *EGR1* is differentially expressed in so many different cancers this was not an unexpected result. What is surprising however is that there appears to be two different populations of patients, with respect to *EGR1* expression. Thirteen of the CRC patients show over-expression of *EGR1* in the tumour (group H), whereas 10 patients show decreased expression of *EGR1* in the tumour (group L). We have shown that the patients that have over-expression of *EGR1* have relatively low levels of *EGR1* in the normal mucosa, compared with the expression of *EGR1* in healthy controls. Conversely the patients that have decreased expression of *EGR1* in the tumour, have relatively high levels of *EGR1*, on a par with the *EGR1* expression levels seen the healthy control population. The paper described by Hong et al 2007, described an over-expression of *EGR1* in the normal mucosa from CRC patients compared with healthy controls which we do not observe in this study. The group with high levels of *EGR1* in the normal mucosa do not have higher *EGR1* levels than the healthy controls. It should be kept in mind however that the study mentioned above was conducted on patients with early onset CRC and from a Chinese population. Unfortunately the patient set used here were anonymous and therefore there is no patient data available for our group of patients, however they are unlikely to be early onset in general as they were collected from the Western General Hospital surgical list, and this possible age difference in study set could explain the differences seen in the two studies.

However, the observance of two different populations with regards their *EGR1* expression in this study is consistent with published data and our Oncomine analysis of *EGR1* expression in normal tissue and tumour of colorectal ancer patients. These studies showed that there is differential expression of *EGR1* in the different stages of tumour development. The first study (Habermann et al, 2007), observed significantly lower levels of *EGR1* in adenoma samples compared with normal tissue, with a

significant increase in *EGR1* from adenoma to carcinoma. As Oncomine analysis showed that lower *EGR1* expression was detected in colorectal adenoma than in normal tissue (Hong et al, 2010), with higher expression of EGR1 in colorectal carcinoma than in normal tissue (Skrzypcak et al) These results could account for why we see two difference groups of *EGR1* expression between normal and tumour in our matched samples.

Given the differential expression seen in the CRC patients we then investigated whether there is any association between the three *EGR1* variants that were determined to have an association with disease phenotype from the case control study conducted by the CCGG. We do not see any correlation with the *EGR1* variants and the differential expression of *EGR1*, however the sample size in this study was small and a study conducted with much larger numbers would be required to conclusively prove or disprove any possible correlation of *EGR1* expression with these variants. However, it is interesting that one group, group L, showed a lack of homozygous wild-type alleles. This group had the highest level of EGR1 in expression in the normal mucosa, with expression levels comparable to the healthy controls. It should also be noted that there was only one homozygous wild-type genotype present in the healthy control group, which showed the highest levels of the EGR1 in the study. It would be of interest to conduct a haplotype analysis and to determine if there is a genetic change on this haplotype background influencing expression.

Somatic tumour mutations of *KRAS* and *BRAF* in the patients was also investigated to determine if they could explain the three observed patient groups. We observed *KRAS* mutations in 9/30 patients and *BRAF* mutations in 2/30 patients. The frequency of the *KRAS* and *BRAF* mutations observed in our group of patients agrees with published data (Ilyas, Straub et al. 1999; Davies, Bignell et al. 2002), with all of the mutations being identified in the two most common mutated codons in both *KRAS* and *BRAF*, codon 12 of KRAS and codon 600 in *BRAF*. *BRAF* mutations are associated with both a sporadic onset microsatellite instability (MSI) phenoytpe and with the CpG island methylator (CIMP) phenotype (Weisenberger, Siegmund et al.

2006). *BRAF* mutations have also been shown to be associated with other features of sporadic MSI-H cancers including advanced age at diagnosis and female sex. *KRAS* mutations are rarely observed in MSI-H phenotype cancers (Kambara, Simms et al. 2004) but are associated with the CIMP phenotype (Toyota, Ahuja et al. 1999). As mentioned we observed that the patients with a *KRAS* or *BRAF* mutation have relatively low levels of *EGR1* expression in the normal mucosa, with 6/11 patients with a mutation showing increased expression of *EGR1* in the tumour. This is very interesting given that the expression of EGR1 is mediated via the MAPK signalling pathway (Chapter 1.3.2, Figure 1.9).

We investigated whether the differential expression of *EGR1* is associated with aberrant methylation of the *EGR1* promoter. A number of promoters regulated by *EGR1* have been found to undergo altered methylation such as the p73 promoter, and the *MDR1* promoter in prostate cancer (Enokida, Shiina et al. 2004; Pipaon, Real et al. 2005). It has been found that the heparanase promoter is CpG island hypomethylated in both bladder and prostate cancer (Ogishima, Shiina et al. 2005) and the expression of heparanase is known to be regulated by *EGR1* in tumour cells (de Mestre, Rao et al. 2005). *EGR1* has several CpG islands in its 5' region (Li and Dahiya 2002) and it has been suggested that CpG methylation may influence the transcription and expression of *EGR1* (Seyfert, McMahon et al. 1990) which made it a reasonable mechanism to be involved in the differential expression of *EGR1* in these patients. However, here there is no evidence to suggest that there is any methylation of the *EGR1* promoter in the CRC cell lines, or any of the CRC patients, either in the normal mucosa or in the tumour in the region of the *EGR1* promoter that we investigated. There was also no evidence of methylation of the *EGR1* promoter in any of the patients in the IBD study. Therefore it can be assumed that aberrant methylation of *EGR1* does not occur in the mucosa of CRC or IBD patients at this region. However it is possible that the expression of *EGR1* is being regulated in a different manner such as the deacetylation of histones, which is mediated by HDACs (Jones and Baylin 2002). There is evidence to show that changes in histone acetylation has an important role in the regulation of *EGR1* and that it is necessary to have a repression chromatin structure in order to prevent the constitutive activation

of *EGR1* (Tur, Georgieva et al. 2010). It would be interesting to further investigate chromatin modifications in the regulation of *EGR1* in CRC patients.

As mentioned previously one of the main mechanisms by which the expression of *EGR1* is regulated is by means of its two co-repressors NAB1 and NAB2, which prevent the constitutive activation of the EGR1 target genes. The NAB proteins regulate the activity of EGR1 by binding to the RI domain of EGR1 and form a complex that binds to the EGR1 binding site. This does not prevent EGR1 from binding to DNA but does inhibit the activation of the target genes (Swirnoff, Apel et al. 1998), EGR1 has been shown to active NAB2 transcription, which can then be repressed by the NAB proteins and appears to act in a negative feedback loop (Kumbrink, Gerlinger et al. 2005). However there are conflicting reports that indicate that NAB2 can act as a positive activator of gene transcription as it has been shown that NAB2 can enhance the EGR1 transcription of IL-2 (Collins, Wolfraim et al. 2006). NAB2 has been shown to be over-expressed in a variety of melanomas (Kirsch, Korradi et al. 1996) and down-regulation of NAB2 is seen in advance prostate cancer (Abdulkadir, Carbone et al. 2001). The NAB proteins could potentially have a role in the regulation of the differential expression see in the CRC patients. Unfortunately there was insufficient RNA available to investigate this in the CRC patients.

One other important investigation of these patient samples would be to look at whether there is differential expression of the EGR1 protein to the same extent as *EGR1* mRNA. It would also be important to determine whether there is any effect on the activation of the downstream targets of *EGR1* such as PTEN and TF. *EGR1* is known to regulate the expression of a variety of genes including several tumour suppressor genes such as *TGFβ1, PTEN, TP53* and *fibronectin* (Baron, Adamson et al. 2006) and the differential expression of *EGR1* may have an impact on its transcriptional activity and ultimately the expression of its downstream target genes. Given the absence of patient data for these patients an analysis into the mutation status of some of these key tumour suppressor genes, such as p53 and PTEN, would be very informative. Similarly it would be possible to determine if these patients are

MSI or MSS and given that MSI and CIN are usually mutually exclusive, this would allow us to determine they type of genetic instatility that exists in these patients and whether it correlates in any way to the differential EGR1 gene expression that we have observed.

Analysis of a group of healthy controls and IBD patients also demonstrated differential expression *EGR1*. The expression of *EGR1* was found to be decreased in the un-inflamed tissue of ulcerative colitis and to a significant level in Crohn's disease. As EGR1 is known to be involved in the regulation of several different genes involved in inflammation it is interesting that the expression of *EGR1* is decreased in un-inflamed mucosa in IBD especially as it has been observed that EGR1 is over-expressed in the inflamed mucosa of both Crohn's disease and ulcerative colitis patients (Subbaramaiah, Yoshimatsu et al. 2004). It was possible to investigate the hypothesis that expression of the *NAB2* repressor protein plays a role in the differential expression of *EGR1* in the healthy controls and IBD patients, however there does not appear to be any over-expression of *NAB2* that may account for lower levels of *EGR1* in these patients.

In order to determine how *EGR1* responds under inflammatory conditions in IBD patients, biospies were cultured overnight in several different inflammatory mediators, LPS, MPD, PGN and TNF. Both LPS and TNF mediators have published links with *EGR1* expression and activity. Lipopolysaccharide (LPS) is a component of the outer membrane of gram-negative bacteria and it induces inflammatory response through the TLR4 receptor and is known to induce *EGR1* expression in many different cell types including monocytes, endothelial cells and macrophages (Coleman, Bartiss et al. 1992; Yao, Mackman et al. 1997; Kadl, Huber et al. 2002; Luyendyk, Schabbauer et al. 2008). It has been shown that EGR1 mediates the LPS induced activation of TNF-α in monocytes (Yao, Mackman et al. 1997). TNF is pro-inflammatory cytokine that is secreted upon activation and has been shown to have increased expression in IBD (Bosani, Ardizzone et al. 2009; Roda, Sartini et al. 2010). *EGR1* was induced by treatment with TNF-α in human fibroblasts and osteosarcoma cells (Cao, Guy et al. 1992; Granet and Miossec 2004). Muramyl

dipeptide (MDP) is the major component of peptidoglycan (PGN) and is found in both gram-positive and gram-negative bacteria. MDP induces and inflammatory response through the activation of NOD2, which activates the NF-κB and IL-12 production. PGN is recognised by the TLR2 receptor. There is little known about the effect of treatment with MDP and PGN of the expression of *EGR1*.

EGR1 is an early response gene and in most cell types the activation of *EGR1* expression is usually seen by 30min, with a return to basal levels by 4hr. However preliminary experiments conducted by the GI unit indicted that *EGR1* expression is induced up to 24hrs in SW480 cells after treatment with all of the inflammatory mediators and this was replicated in my own data. It is known that the induction of EGR1 by different stress stimuli can have opposing effects depending on the stimuli and cell type (Yu, de Belle et al. 2004). Interestingly in endometrial carcinoma cells EGR1 protein expression was shown to be induced up to 24hrs after treatment with ER-α, and constitutive activation of the EGFR/ERK1/2 pathway can result in constant high levels of EGR1 protein in prostate cells, indicating that the induction of *EGR1* can be sustained in some cell types (Saegusa, Hashimura et al. 2008; Sauer, Gitenay et al. 2010).

In the mucosa of the healthy controls *EGR1* expression is significantly induced after LPS and TNF treatment and the majority of patients show induction of *EGR1* after these treatments, indicating that *EGR1* is induced by inflammatory mediators in healthy normal mucosa as expected. Interestingly there is little or no induction of EGR1 expression after MDP or PGN treatment, indeed 35% of patients show a decrease in EGR1 expression after MDP treatment and 57% show a decrease after PGN. It could be that *EGR1* is not activated to the same extent after exposure to MDP or PGN, or that unlike after LPS and TNF, the expression of *EGR1* is more transient in response to these treatments in normal mucosa. It should also be kept in mind that the sample numbers for each patient group in this study are small and larger numbers would be required to get a more accurate picture.

The un-inflamed mucosa in ulcerative colitis patients show significant induction of *EGR1* after treatment with TNF as in the healthy controls. However *EGR1* expression does not seemed to be induced to the same extent after LPS treatment, indeed only 49% of the UC patients show an induction of *EGR1* after LPS. Interestingly the treatment of UC mucosa with PGN actually seems to cause a significant decrease in the levels of *EGR1*. However it is surprising that there is such a difference after treatment with MDP and PGN in these patients, as MDP is a component of PGN.

In the Crohn's disease patients there is a significant decrease in *EGR1* expression after LPS treatment, with 70% of the patients showing lower levels of *EGR1* after treatment with LPS. Although in the overall population analysis only LPS shows a significant decrease, 66% of patients show a decrease after TNF treatment and 66% show a significant decrease after MDP treatment. It is interesting that the Crohn's disease patients show such an aberrant response after MDP treatment. NOD2 is the main susceptibility gene for Crohn's disease and has been shown to be activated by MDP (Strober, Murray et al. 2006). A yeast two-hybrid screen conducted by the GI unit identified EGR1 as a novel binding protein with NOD2, which will be investigated in the next chapter. However it is intriguing that the Crohn's disease patients show such an aberrant response of *EGR1* induction after MDP treatment given this link with NOD2.

It is clear however that in the un-inflamed tissue of the IBD patients there is an aberrant induction of *EGR1* after inflammatory treatments. It would interesting to determine if the protein levels of EGR1 are similarly effected, and whether the aberrant response of *EGR1* to these inflammatory stimuli has an effect on the downstream inflammatory targets of EGR1 such as mPGES-1. It has been suggested that there is an important role for prostaglandin E_2 (PGE_2) in inflammatory bowel disease, and it has been shown that there are increased levels of mPGES-1, which is important for the synthesis of PGE_2, in inflamed intestinal mucosa in both CD and UC (Subbaramaiah, Yoshimatsu et al. 2004). EGR1 is known to regulate the transcription of the mPGES-1 promoter (Naraba, Yokoyama et al. 2002), and it has

been shown that EGR1 is necessary for the TNFα-mediated induction of mPGES-1. Interestingly $mPGE_2$ is over-expressed in CRC (Yoshimatsu, Golijanin et al. 2001). It would therefore be of considerable interest to determine if the EGR1-mediated transcription of mPGES-1 is aberrant in IBD.

In conclusion, we have shown that there is differential expression of colonic *EGR1* in diseased colon, in both CRC and IBD. In CRC there is both increased expression and suppression of *EGR1* in the tumour and it is unknown how this differential expression is regulated. We know that it is unlikely to be associated with aberrant methylation of the EGR1 promoter in normal mucosa or tumour or with the *EGR1* common variants in the germline, although almost half of those patients with lower EGR1 expression in the tumour has a mutation in the RAF signalling pathway. There is also differential expression of *EGR1* between healthy controls and patients with IBD, with the IBD patients also showing an aberrant *EGR1* response to inflammatory mediators. It is unlikely that this differential expression is associated with aberrant methylation of the *EGR1* promoter or with the expression of the EGR1 co-repressor protein NAB2. Future work would be to determine how the differential expression of *EGR1* is regulated in diseased colon and understanding what effect this differential expression has on the downstream targets of *EGR1* in CRC and IBD would be of great interest.

4 Chapter 4

4.1 Introduction

EGR1 is a transcription factor that is known to be localised in the nucleus, but has also been shown to localise in the cytoplasm. EGR1 contains a bipartite nuclear localisation signal located in the zinc-finger domain of EGR1 (Gashler, Swaminathan et al. 1993), and it has been suggested that the components of the cytoskeleton may have a role to play in the nuclear translocation of EGR1 (Mora, Olivier et al. 2004). EGR1 has been shown to localise with microtubules, with both α- and β-tubulin, in benign prostate cells, whereas no interaction is visible in malignant prostate cells (Mora, Olivier et al. 2004). Similarly preliminary results by the CCGG demonstrated that EGR1 co-localises with the cytoskeleton in skin fibroblasts and CRC cells by immunocytochemistry and have demonstrated a potential interaction with γ-tubulin by IP. Therefore further work was required to investigate the potential interaction with cytoskeleton components such as γ-tubulin and α-tubulin in CRC cell lines.

An interesting finding by the Gastrointestinal Unit was that EGR1 is an interacting partner of the IBD susceptibility NOD2 protein by yeast-two hybrid. It was decided to investigate this interaction in colorectal cancer cell lines and ultimately to determine whether or not this interaction was localised in the cytoplasm or nucleus, and whether it was altered upon stimulation of EGR1 by stress stimuli such as LPS. The aim of this chapter was to investigate the cellular role of EGR1 in colorectal disease by advancing preliminary work preformed by the CCGG and GI groups, including specific cellular localisation and interaction with other proteins. The cellular localisation of EGR1 in CRC cells by western blot and immunoctyochemistry is investigated and whether NOD2 localisation coincides with EGR1 in these cells. The interaction between EGR1 and NOD2 is investigated using immuoprecipitation in both untreated and LPS treated cells. The interaction between EGR1 and tubulin components are also confirmed using immuoprecipitation in SW480 cells.

4.2 Methods

4.2.1 Cell Culture

Cell culture was conducted as detailed in the methods section (2.1.2).

4.2.2 siRNA knockdown of EGR1

HRT18 cells were plated in equal amounts into a 6 well tissue culture plate and grown until ~80-90% confluent. siRNA transfections were performed using Lipofectamine 2000 reagent (Invitrogen) in Optimem-1 medium (Gibco) and the Stealth siRNA (Invitrogen) as per manufacturer's instructions. Cells were transfected in antibiotic-free media for 48 hours. Cells were subsequently harvested, the protein was extracted and resolved by Western Blot analysis as detailed in methods (2.8).

Stealth siRNA Primers	
EGR1HSS103117 **(1)**	UCU CCC AGG ACA AUU GAA AUU UGG U AGC AAA UUU CAA UUG UCC UGG GAG A
EGR1HSS103118 **(2)**	GAU CUC UGA CCC GUU CGG AUC CUU U AAA GAA UCC GAA CGG GUC AGA GAU C
EGR1HSS103119 **(3)**	CCA UGG ACA ACU ACC CUA AGC UGG A UCC AGC UUA GGG UAG UUG UCC AUG G
Scrambled control – mutant SOCS-1	CUA UCU AAG UUA CUA CCC CUU TT

Table 4.1: siRNA primers

4.2.3 Treatment of cells

HRT18 and SW480 cells were treated with various concentrations of Curcumin (Sigma Aldrich, UK) ranging from 10nM to 15μM, for time points between 30min to 22 hours. The stock solution of Curcumin was prepared in dimethyl sulfoxide (DMSO) at a concentration of 10mM, and all subsequent concentrations of Curcumin were prepared in DMSO to a final volume of 500μl of media. DMSO only controls were included in all experiments.

HRT18, SW480 and Vaco425 cells were also treated with concentrations of LPS ranging from 100ng/ml to 10μg/ml, for various time lengths between 30min to 22 hours. The stock solution of LPS was prepared in PBS at a concentration of 1mg/ml and all subsequent concentrations of LPS were prepared in PBS.

4.2.4 Preparation of protein extracts

Two methods of nuclear and cytoplasmic extraction were tested as detailed below, with longer extraction times for the cytoplasmic extract in protocol (B).

4.2.4.1 Preparation of Cytoplasmic and Nuclear Extracts (A)

To prepare cytoplasmic and nuclear extracts the cells were resuspended in 100μl of Normal Lysis Buffer, supplemented with protease inhibitors. Following an incubation of 5min on ice, the cells were centrifuged at 13,000 rpm for 20sec. The supernatant, which is the cytoplasmic extract, was removed and transferred to pre-chilled eppendorf tubes. The cell pellet was then resuspended in 30μl Hypotonic Lysis Buffer, supplemented with protease inhibitors (2.8.3.1) and incubated on ice for 30min. The cells were centrifuged at 13,000 rpm for 5min and the supernatant, consisting of the nuclear extract, was removed and transferred to pre-chilled eppendorfs.

4.2.4.2 Preparation of Cytoplasmic and Nuclear Extracts (B)

To prepare cytoplasmic and nuclear extracts the cells were resuspended in 100μl of Normal Lysis Buffer, supplemented with the same protease inhibitors (2.8.3.1). The cells were then centrifuged at 6,000 rpm for 20min. The rest of the protocol continues the same as above from this stage.

Ultimately both extraction methods resulted in the same localisation of EGR1 in the CRC cell lines and the data in this chapter illustrates nuclear and cytoplasmic extracts prepared as detailed in the (A) protocol.

4.2.4.3 Preparation of cellular extracts into more distinct cellular compartments

In order to separate the extracts into four cellular compartments (cytoplasmic, membrane/organelle, nuclear and cytoskeleton) the cells were grown until confluent in a T25cm^3 flask. The cells were then extracted using the ProteoExtract Subcellular Proteome Extraction Kit (Calbiochem) as per the manufacturer's instructions.

4.2.5 Western Blotting

Protein extracts were resolved by western blot analysis as detailed in the methods section (2.8.4).

4.2.6 Antibodies

The following antibodies were used in this chapter:

Antibody	Manufacturer	Dilution	Incubation conditions
EGR1 Goat mAB	R&D Systems AF2818	1:1000 (WB) 1:50 (IC)	4°C - overnight
EGR1 (588) Rabbit polyclonal	Santa Cruz Biotechnology sc-110	1:1000 (WB)	4°C - overnight
EGR1 (C-19) Rabbit polyclonal	Santa Cruz Biotechnology sc-189	1:1000 (WB)	4°C - overnight
NOD2 Mouse mAB	Cayman Chemicals Cat # 10004942	1:200 (WB) 1:200 (IC)	Room temperture. – 1hour
γ-tubulin Mouse mAB	Sigma Chemical T6557	1:5000 (WB) 1:200 (IC)	Room temp. – 1hour
α-tubulin Mouse mAB	Sigma Chemical T9026	1:1000 (WB) 1:200 (IC)	Room temp. – 1hour
NAB2 D-7 Mouse mAB	Santa Cruz Biotechnology sc-48416	1:1000 (WB)	Room temp. – 1hour
β-actin IgM Mouse	Sigma Chemical	1:400 (WB)	Room temp. – 1hour
c-Myc Mouse mAB	Santa Cruz Biotechnology sc-40	1:200 (WB)	Room temp. – 1hour
Secondary Antibodies			
Goat anti-Mouse IgG-HRP	Santa Cruz Biotechnology (Sc-2005)	1:2000 (WB)	Standard
Donkey anti-Goat IgG-HRP	Santa Cruz Biotechnology (Sc-2056)	1:2500 (WB)	Standard

Goat Anti-Rabbit IgG - Horseradish Peroxidase	Amersham Biosciences (RPN4301)	1:2000 (IC)	Standard
Sheep Anti-Mouse IgG - Horseradish Peroxidase	Amersham Biosciences (RPN4201)	1:2000 (IC)	Standard

Table 4.2: Primary and secondary antibodies

WB indicates dilution used in Western blotting techniques, IC indicates dilution used for immunocytochemistry.

4.2.7 Immunoprecipitation

SW480 cells were seeded into a T75cm3 flask and grown until 80-90% confluent. The cells were harvested and the proteins were extracted in 500µl of lysis IP buffer with protease inhibitors added. The protein concentration was determined by Bradford assay. The agarose G beads where blocked in lysis buffer containing BSA (100µg/ml tRNA, 100µg/ml BSA, 100µ/ glycogen). The agarose G beads were washed 3 times in IP buffer and the protein extracts were pre-cleared in washed agarose G beads for 1 hour at 4 ^{o}C. 500µg of protein lysate was incubated in 2µg of antibody overnight at 4 ^{o}C while rotating. 15µl of pre-washed agarose G beads was added and incubated at 4 ^{o}C for 1 hour while rotating. The beads were washed and 50µl of 2 times running buffer was added. The samples were boiled for 5 min, and centrifuged at 14,000rpm at room temperature for 5min. The supernatant was separated on a 10% resolving gel and visualised by Western Blot analysis.

	Salt	Detergent	Buffer
IP lysis buffer 1	150mM NaCl	0.5% NP40	50mM Hepes (pH 7.4)
IP lysis buffer 2	150mM NaCl	0.5% Tween	50mM Hepes (pH 7.4)
IP lysis buffer 3	300mM NaCl	0.5% Tween	50mM Hepes (pH 7.4)
IP lysis buffer 4	300mM NaCl	1% Triton x100	50mM Hepes (pH 7.4)
IP lysis buffer 5	400mM KCl	1% NP40	25mM Tris

Table 4.3: Differences between the IP lysis buffers tested

4.2.8 Immunocytochemistry

Initially two methods of immunocytochemistry were tested.

4.2.8.1 Immunocytochemistry method using methanol:acetone

Cells were spilt and 500µl of cell suspension was plated into a 6 well plate containing a sterile coverslip and grown overnight at 37^{o}C. The medium was removed and cells/coverslip washed in 5ml cold PBS. The cells were fixed with 3ml of methanol:acetone (1:1) and incubated at –20^{o}C for at least 30min. After removal of the methanol:acetone the cells were washed with PBS while shaking for 10mins, 3 times. A blocking solution of 10% donkey serum in PBS was added to the cells for 30min at room temperature. After removal of the donkey serum, the cells were incubated with primary antibody in 10% donkey serum and incubated at room temperature for 1 hour. After washing with 0.15% Tween-10/PB for 10min while shaking, the cells were incubated with a secondary antibody in 1.5% donkey serum for 30min at room temperature in the dark due to the use of fluorescent secondary antibody. The cells were washed 3 times for 10min as before and the cover slips were placed cell side down onto microscope slides containing a few drops of DAPI-Vectashield. The slides were stored in the dark at 4^{o}C and the images were visualised using fluorescence microscopy.

4.2.8.2 Immunocytochemistry method using formaldehyde

Cells were spilt and plated as above. The cells were fixed with 1ml of 4% formaldehyde and incubated at room temperature for 20min. After removal of the formaldehyde, the cells were rinsed several times with PBS, and a solution of 0.5% Triton-x was added to the cells for 2min at room temperature. The cells were washed with PBS 3 times for 5min while shaking and then the cells were blocked with 5% BSA in PBS all day. The cells were then washed with PBS 3 times for 5min as before and incubated with a primary antibody in 5% BSA in PBS overnight at 4^{o}C. After washing with PBS 3 times for 5min while shaking, the cells were incubated with a secondary antibody in 5% BSA for 1hr at room temperature in the dark. The cells were washed 3 times for 10min as before and the cover slips were placed cell side down onto microscope slides containing a few drops of DAPI-Vectashield. The slides were stored in the dark at 4^{o}C and the images were visualised using fluorescence microscopy.

The immunocytochemistry method using formaldehyde to fix the cells gave much better immunostaining and consistent results. All of the immunofluorescence experiments detailed in this chapter were conducted using the formaldehyde protocol.

4.3 Results

4.3.1 Determination of accurate tools for EGR1 protein analysis

A number of different antibodies for EGR1 are commercially available and had been used with mixed results previously in the lab. Whole cell extracts of four colorectal cancer cell lines (HRT18, HCT116, SW480 and Vaco425) were prepared, resolved on a Western blot and probed with three different commercially available antibodies. Two antibodies that had been used previously in the lab are EGR1 anti-rabbit available from Santa Cruz Biotechnology (Table 4.2). A third antibody tested was EGR1 anti-goat from R&D Systems (AF2818). Figure 4.1 illustrates the membrane probed with the three different antibodies. Antibody A (Figure 4.1 A) is EGR1 α-rabbit (sc-110), however EGR1 expression is barely evident in any of the three cell lines. Using antibody B (EGR1 α-rabbit, Santa Cruz Biotechnology, Sc-189) demonstrates expression in the cell lines, however in HCT116 and Vaco425 multiple bands are evident, and only a weak band is evident in HRT18 and SW480 cells (Figure 4.1 B). The final EGR1 antibody (EGR1 α-goat, R&D Systems) gave a clear single band in HCT116 and SW480 cells, and a weak band evident in HRT18 cells located at the correct position for EGR1 at ~ 80kDa (Figure 4.1 C). Therefore it was decided to use the third antibody for all future experiments. After further optimisation it was determined that for best results the EGR1 antibody requires overnight incubation in milk PBS/Tween along with washing the membrane for 1 hour 3 times after incubation with both the EGR1 antibody and the secondary anti-goat antibody.

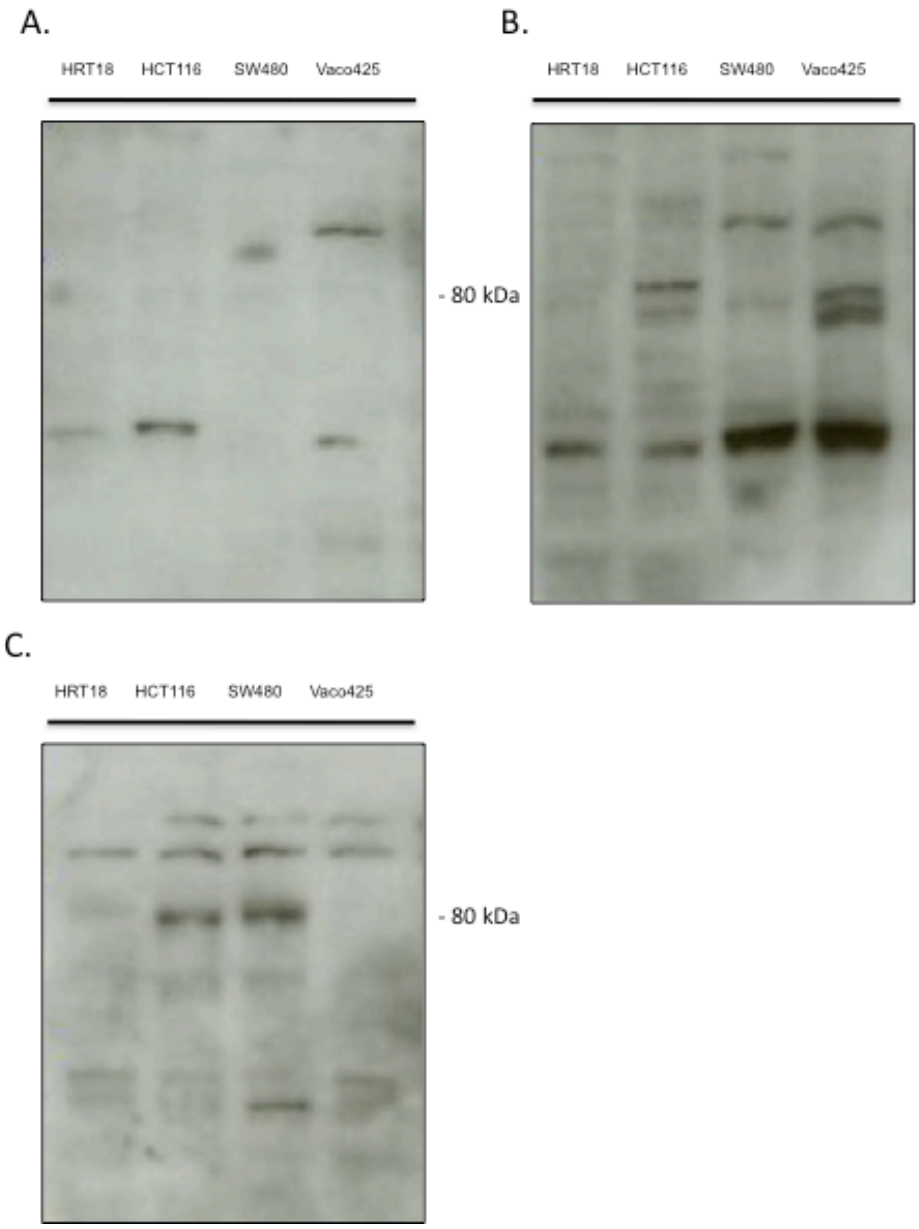


Figure 4.1: EGR1 antibody in CRC cell lines

Whole cell extracts were prepared from HRT18, HCT116, SW480 and Vaco425, resolved on a 10% acrylamide gel and probed with three different anti-EGR1 antibodies **A.** The anti-EGR1 antibody (Santa Cruz Biotehcnology, sc-110) gave a weak band of EGR1 expression in all three cell lines. **B.** The anti-EGR1 (Santa Cruz Biotechnology, sc-189) shows a stronger a band but there appears to be multiple bands. **C.** The third anti-EGR1 (R&D systems, AF2818) gave a clear single band for EGR1 expression at the correct size of ~80kDa.

In order to confirm the presence and the detection of EGR1 by the antibody chosen, the expression of EGR1 was knocked down in HRT18 cells using three different siRNA primers for 48 hours. Whole cell extracts were prepared, resolved on a Western blot and probed with the anti-EGR1 antibody. The membranes were subsequently probed with an anti-actin antibody to verify protein loading. It is clear from figure 4.2 that both siRNA primer 1 and 2 reduces EGR1 expression after 48 hours. This confirmed that the antibody being used for analysis is detecting the EGR1 protein and it is the correct size of ~80 kDa.

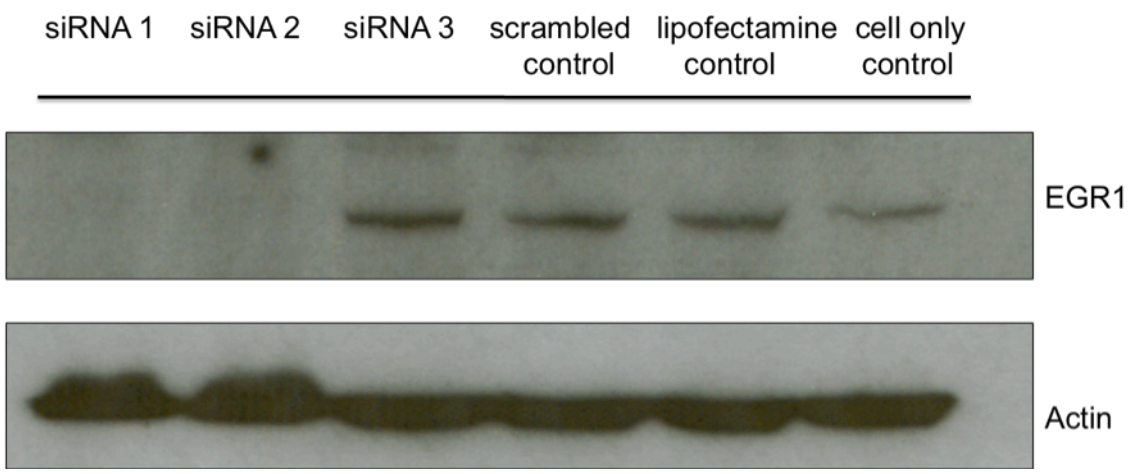


Figure 4.2: Knockdown of EGR1

HRT18 cells were transfected with three different siRNA primers for 48 hours. Cells were also transfected with a scrambled control and a lipofectamine only control. Whole cell extracts were prepared (30μg) and resolved on a 10% acrylamide gel. The membranes were probed with anti-EGR1 and anti-actin as a loading control.

This clearly demonstrated that the EGR1 antibody chosen detects a band of correct size for EGR1, with EGR1 siRNA reducing this expression. However we wanted to ensure that the EGR1 protein being detected responds to induction by stress stimuli. As it has been shown that different cell lines respond differentially to treatment with stress stimuli the experiment was conducted using both HRT18, which has low levels of EGR1 expression, and SW480 cells which has higher levels of EGR1 at both the mRNA and a protein level (Figure 3.1; Section 3.3.1).

SW480 (A) and HRT18 (B) cells were treated with various concentrations of curcumin, ranging from 100nM to 15μM, for a 3 hour time period and using a concentration of 1μM of curcumin over a time period ranging from 30 minutes to 21 hours (Figure 4.3 A, B). There is an increase in the levels of EGR1 protein in both cell lines, after treatment of 500nM of curcumin or higher. In SW480 cells, there is an increase at 1μM, taking into account differences in protein loading and the induction of EGR1 with 1μM of curcumin appears to occur after 30 minutes to 6 hours (Figure 4.3 A). In HRT18 cells the protein levels are increased at concentrations 500nM to 10μM with EGR1 levels increasing after treatment with 1μM between 1 hour and 3 hours (Figure 4.3 B).

The experiment was repeated using LPS to induce expression of EGR1. The cells were treated with concentrations of LPS ranging from 100ng/ml to 10μg/ml for 6 hours, and treated with 1μg/ml over a time period, ranging from 30min to 21 hours (Figure 4.3 C, D). In SW480 cells there is an increase in EGR1 protein levels as the concentration of LPS increases, and an increase in EGR1 after 1 hour with 1μg/ml LPS (Figure 4.3 C). There is no clear increase of EGR1 in HRT18 at any particular concentration, but an increase in EGR1 protein levels after 1 hour with 1μg/ml treatment LPS (Figure 4.3 D). This experiment demonstrates that there is induction of EGR1 evident using the α-EGR1 antibody at the correct band size of ~80kDa using both curcumin and LPS in both HRT18 and SW480 cells.

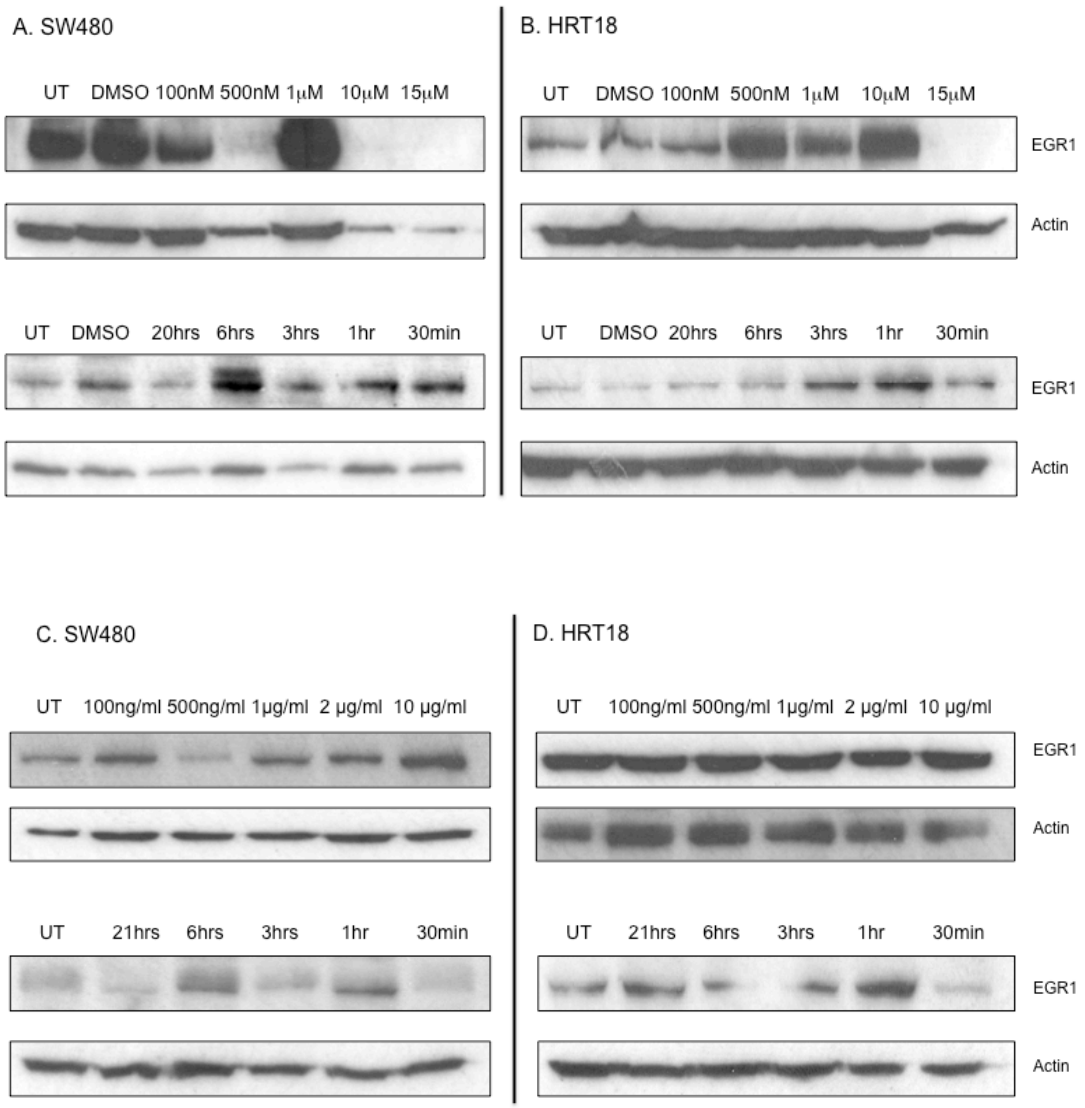


Figure 4.3: CRC cells treated with Curcumin and LPS

Cells were treated with curcumin and LPS over a range of concentrations and over a time period as detailed below. Whole cell extracts were prepared (30ng) and probed with anti-EGR1 and anti-actin. **A.** SW480 cells and **B.** HRT18 cells were treated with various concentrations of curcumin (100nM, 500nM, 1μM, 10μM and 15μM) for 3 hours and treated with 1μM curcumin for over a time period (20 hrs, 6 hrs, 3 hrs, 1hr and 30 min). **C.** SW480 cells and **D.** HRT18 cells were treated with various concentrations of LPS (100ng/ml, 500ng/ml, 1μg/ml, 2μg/ml and 10μg/ml) for 6 hours and then treated with 1μg/ml of LPS over a time course (21 hrs, 6 hrs, 3 hrs, 1hr and 30 min).

4.3.2 Expression of EGR1 in CRC cells

In order to determine the protein levels of EGR1 in various colorectal cancer cell lines, whole cell extracts of HRT18, HCT116, SW480 and Vaco425 were prepared with 30μg of protein, resolved on an acrylamide gel and probed with the anti-EGR1 antibody (Figure 4.4). High protein levels of EGR1 are observed in both HCT116 and SW480 cells, but lower protein levels in HRT18 and Vaco425. Interestingly Vaco425 has the highest relative *EGR1* mRNA expression levels in the qRT-PCR analysis from the previous chapter but shows the least amount of protein in this case. The next investigation was to determine the localisation of the EGR1 protein in CRC cells, hence cytoplasmic and nuclear protein extracts of the four CRC cell lines above were prepared at a concentration of 30μg of protein. This again shows lower levels of EGR1 in HRT18 and Vaco425 cells. It appears that EGR1 is localised in the nucleus and the protein is also detected strongly at times in the cytoplasm in all cell lines tested except for Vaco425. Indeed EGR1 is possibly increased in the cytoplasm of HRT18 cells rather than the nucleus, which is intriguing as the HRT18 cell line contains the wild-type allele for all three *EGR1* SNPs.

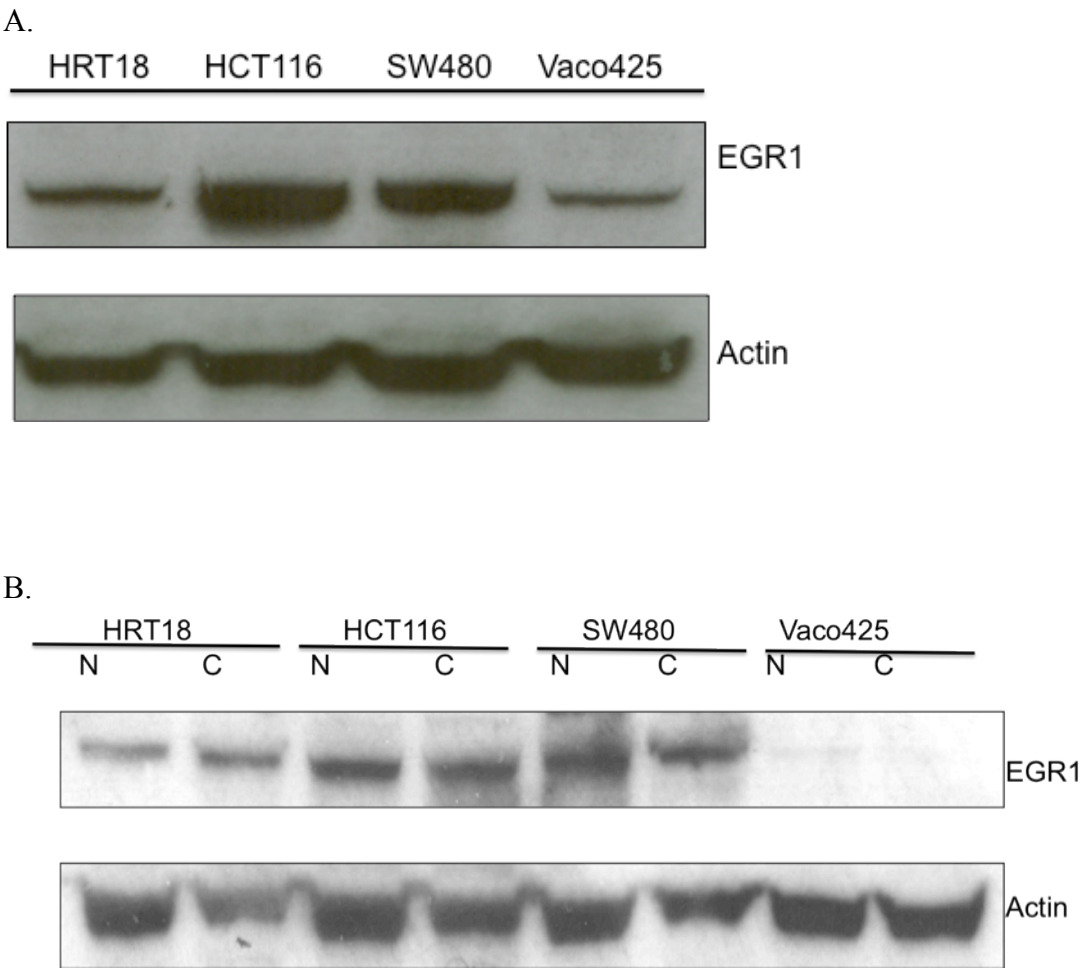


Figure 4.4: Expression of EGR1 in CRC cell lines

Protein extracts (30μg) were prepared from HRT18, HCT116, SW480 and Vaco425, resolved on a 10% acrylamide gel and probed with anti-EGR1 antibody and stripped and re-probed with anti-actin antibody as a loading control. **A.** Illustrates EGR1 protein levels in whole cell extracts **B.** EGR1 levels in nuclear (N) and cytoplasmic (C) extracts.

Immunofluorescence experiments were conducted to visualise the localisation of the EGR1 protein in HRT18 and SW480 cells. It was decided to further investigate the expression and localisation of EGR1 in these two cell lines specifically, as they have different genotypes of the three EGR1 SNPs that were investigated in the previous chapter. HRT18 is homozygous wild type for the three EGR1 SNPs and SW480 is homozygous variant.

As figure 4.5 illustrates, EGR1 is localised in both the nucleus and the cytoplasm in both these cells, with stronger cytoplasmic staining evident in the HRT18 cells, in agreement with the previous analysis.

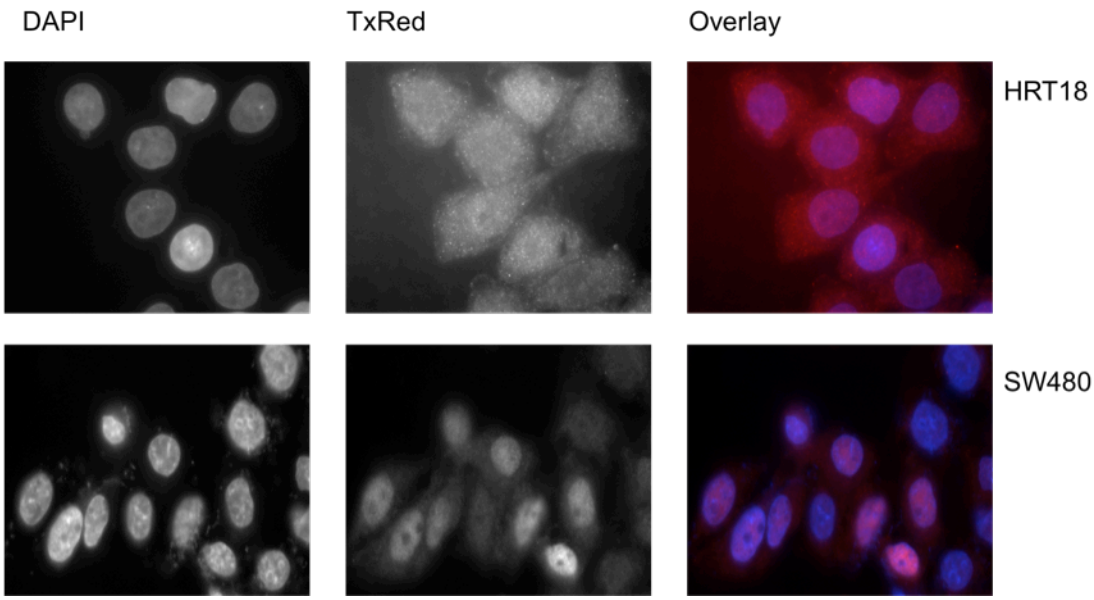


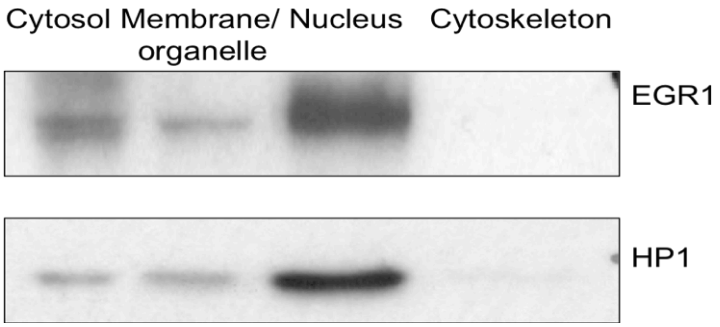
Figure 4.5: Localisation of EGR1 in CRC cell lines using immunofluorescence

HRT18 and SW480 cells were grown on cover slips and fixed with formaldehyde. The cells were immunostained with EGR1 anti-goat and TxRed-conjugated Goat Anti-Rabbit IgG - Horseradish Peroxidase, with DAPI to stain the nuclei (magnification x40).

As EGR1 is a transcription factor, it was expected that the protein would be predominantly localised in the nucleus. However, as discussed there is also evidence to suggest that EGR1 may shuttle between the cytoplasm and the nucleus and so a subcellular localisation kit was used to further determine the exact localisation of EGR1. The protein extract can be separated into four subcellular fractions, the cytosol fraction, a membrane bound/organelle fraction, the nuclear fraction and the cytoskeleton fraction. Figure 4.6 A illustrates that, as expected in the SW480 cells, EGR1 is predominantly localised in the nucleus with protein evident in the cytosol and membrane bound/organelle fraction also. The membrane was subsequently probed with anti-HP1 as a nuclear control to determine correct separation of the protein fractions. There is a strong band in the nuclear fraction for HP1 as expected, however there are also faint bands present in the cytoplasmic and organelle/membrance bound fraction, indicating a slight contmination of nuclear protein the the cytoplasmic fractions. Figure 4.6 B shows the localisation of EGR1 over a 16-hour time period.

Due to inconsistency and low levels of EGR1 obtained from HRT18 cells when using the subcellular localisation kit it was not possible to conduct all experiments in this cell line as hoped. Similarly low levels of NOD2 is also observed in HRT18 cells (as will be demonstrated in Figure 4.7) so SW480 will used in the majority of the future experiments.

A.



B.

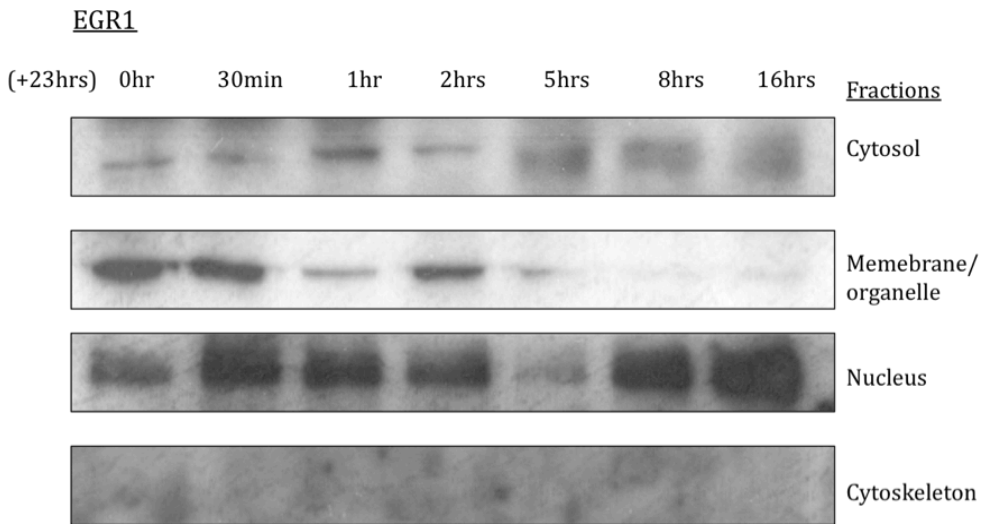


Figure 4.6: Localisation of EGR1 in cellular sub-compartments

A. SW40 cells were extracted into four cellular sub-compartments, the cytosol, membrane bound/organelle, nuclear and cytoskeleton and 30μg of protein was resolved on a 10% gel and probed with anti-EGR1 antibody. The membrane was stripped then probed with anti-HP1 as a positive control for nuclear protein. **B**. SW480 cells were extracted into the four cellular compartments as above at several different time points, starting at time point 0hr (23 hours after splitting). 30μg of protein was resolved on a 10% gel and probed for EGR1.

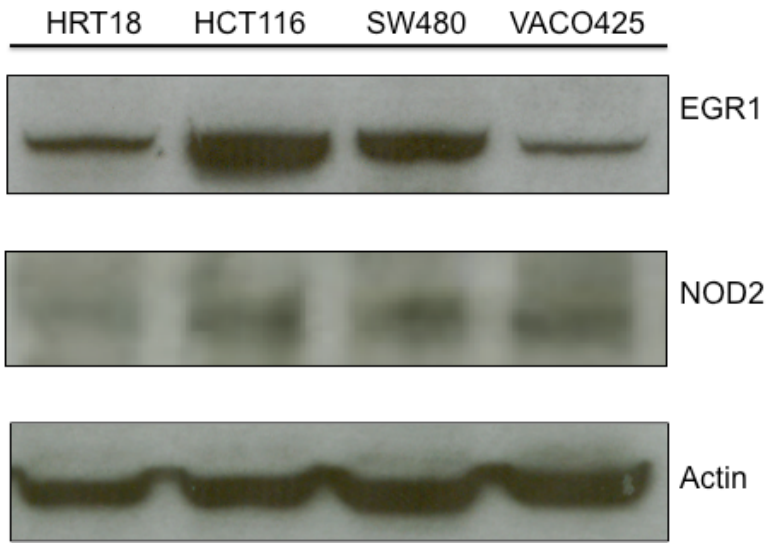
4.3.3 Investigation into a potential interaction of EGR1 and NOD2

4.3.3.1 Localisation and expression of NOD2 in CRC cells

In order to confirm a potential interaction between EGR1 and NOD2, we first wanted to characterise the expression and localisation of NOD2 in CRC cell lines and determine if NOD2 localisation coincides with EGR1. Whole cell extracts were prepared from HRT18, HCT116, SW480 and Vaco425 cells, 30μg of protein were resolved on an acrylamide gel and probed with the anti-NOD2 antibody (Figure 4.7 A) The protein levels of EGR1 as demonstrated in Figure 4.1.2 are included in the figure also. There are low levels of NOD2 expressed in HRT18 cells, with higher levels evident in HCT116, SW480 and Vaco425 cells.

Nuclear and cytoplasmic extracts from the same four cell lines were used to determine the cellular localisation of NOD2 (Figure 4.7 B). As expected from the literature the majority of NOD2 is localised in the cytoplasm in HRT18, HCT116 and SW480 cell lines. However in the Vaco425 cells, NOD2 is detected equally in the nuclear extract as well as the cytoplasmic extract.

A.



B.

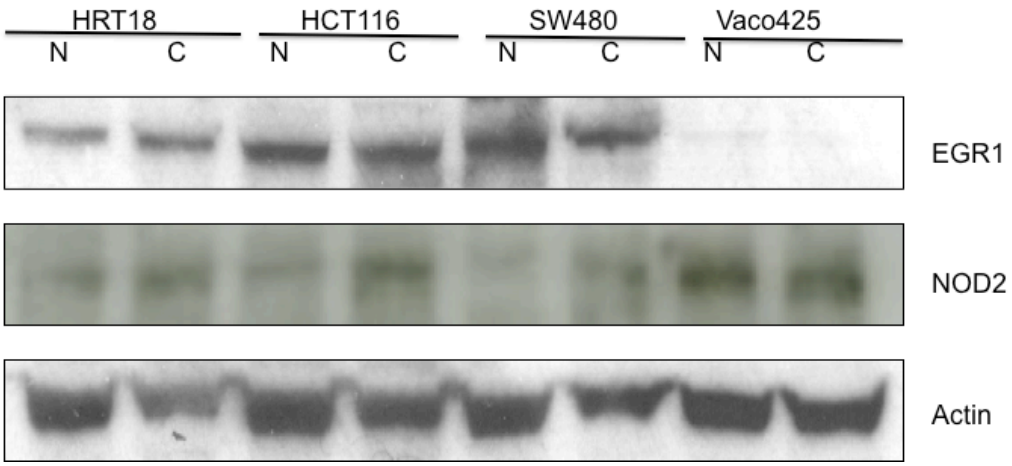
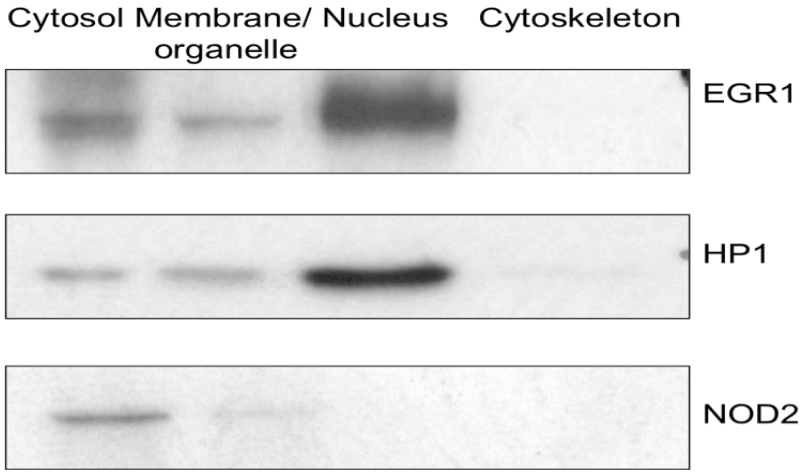


Figure 4.7: Expression of NOD2 in CRC cell lines

Protein extracts (30μg) were prepared from HRT18, HCT116, SW480 and Vaco425, resolved on a 10% acrylamide gel and probed with anti-EGR1 antibody, anti-NOD2 and anti-actin antibody as a loading control. The membranes were stripped and re-probed **A.** Illustrates EGR1 and NOD2 protein levels in whole cell extracts **B.** EGR1 and NOD2 levels in nuclear (N) and cytoplasmic (C) extracts.

There has been evidence published to suggest that NOD2 localises to the membrane in intestinal epithelial cells and this localisation is necessary for the MDP induction of NOD2-mediated activation of NF-κB (Barnich, Aguirre et al. 2005). Therefore a subcellular localisation experiment was used to determine if NOD2 is localised in any other cellular compartment in the SW480 cells. The protein extracts were separated into four subcellular fractions, the cytosol fraction, a membrane bound/organelle fraction, the nuclear fraction and the cytoskeleton fraction and probed with anti-NOD2 antibody. The membrane was also probed with EGR1 to determine if EGR1 and NOD2 may be localised in the same compartment (Figure 4.8 A). Again it is evident that NOD2 is localised in the cytoplasm in SW480 CRC cells, with some protein detected in the membrane bound/organelle fraction also. Immunofluorescence shows that NOD2 is localised in cytoplasm in both HRT18 and SW480 cells (Figure 4.8 B). The cells were stained with FITC labelled α-mouse secondary antibody to visualise NOD2.

A.



B.

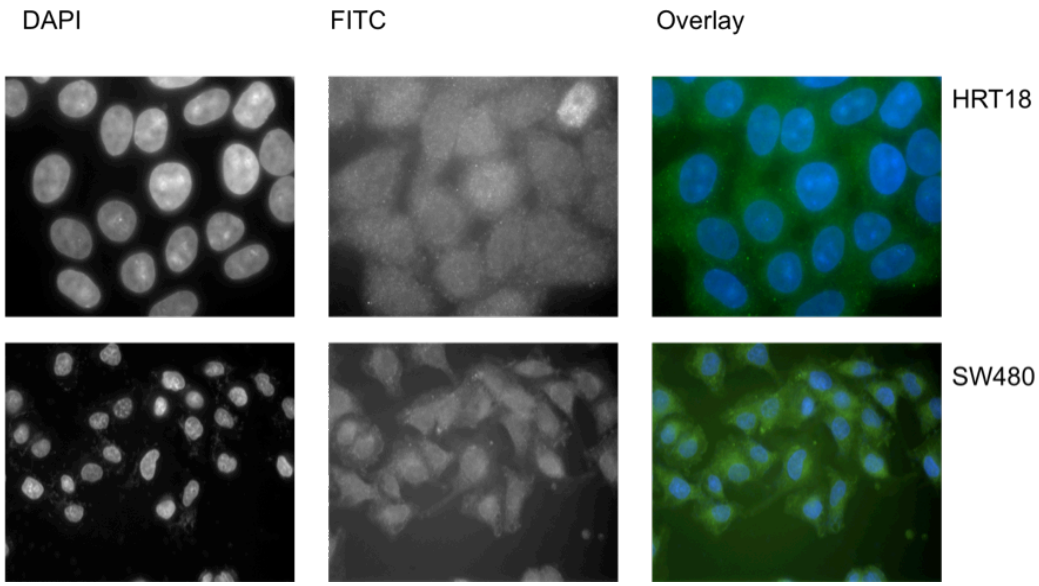


Figure 4.8: Localisation of NOD2 in CRC cell lines

A. SW40 cells were extracted into four cellular compartments, the cytosol, membrane bound/organelle, nuclear and cytoskeleton and 30μg of protein was resolved on a 10% gel and probed with anti-NOD2 and anti-EGR1 antibody. The membrane was then probed with anti-HP1 as a positive control for nuclear protein**.** The membrane was stripped and re-probed with all antibodies. **B.** HRT18 and SW480 cells were grown on cover slips and fixed with formaldehyde. The cells were immunostained with NOD2 anti-mouse and FITC-conjugated Mouse AntiSheep IgG - Horseradish Peroxidase, with DAPI to stain the nuclei (magnification x40).

4.3.3.2 EGR1 and NOD2 co-localises in CRC cells

Although EGR1 is predominantly localised in the nucleus, it is evident from the previous results and the literature that EGR1 can localise in the cytoplasm, so further immunocytochemistry experiments were conducted to determine if EGR1 and NOD2 co-localise in CRC cells. SW480 and HRT18 cells were grown on cover slips and fixed using formaldehyde. The cells were co-stained with FITC-conjugated NOD2 (1:200) and TxRed-conjugated EGR1 (1:50). As figure 4.9 illustrates, it does appear that EGR1 and NOD2 can co-localise in the cytoplasm in both HRT18 and SW480 cells.

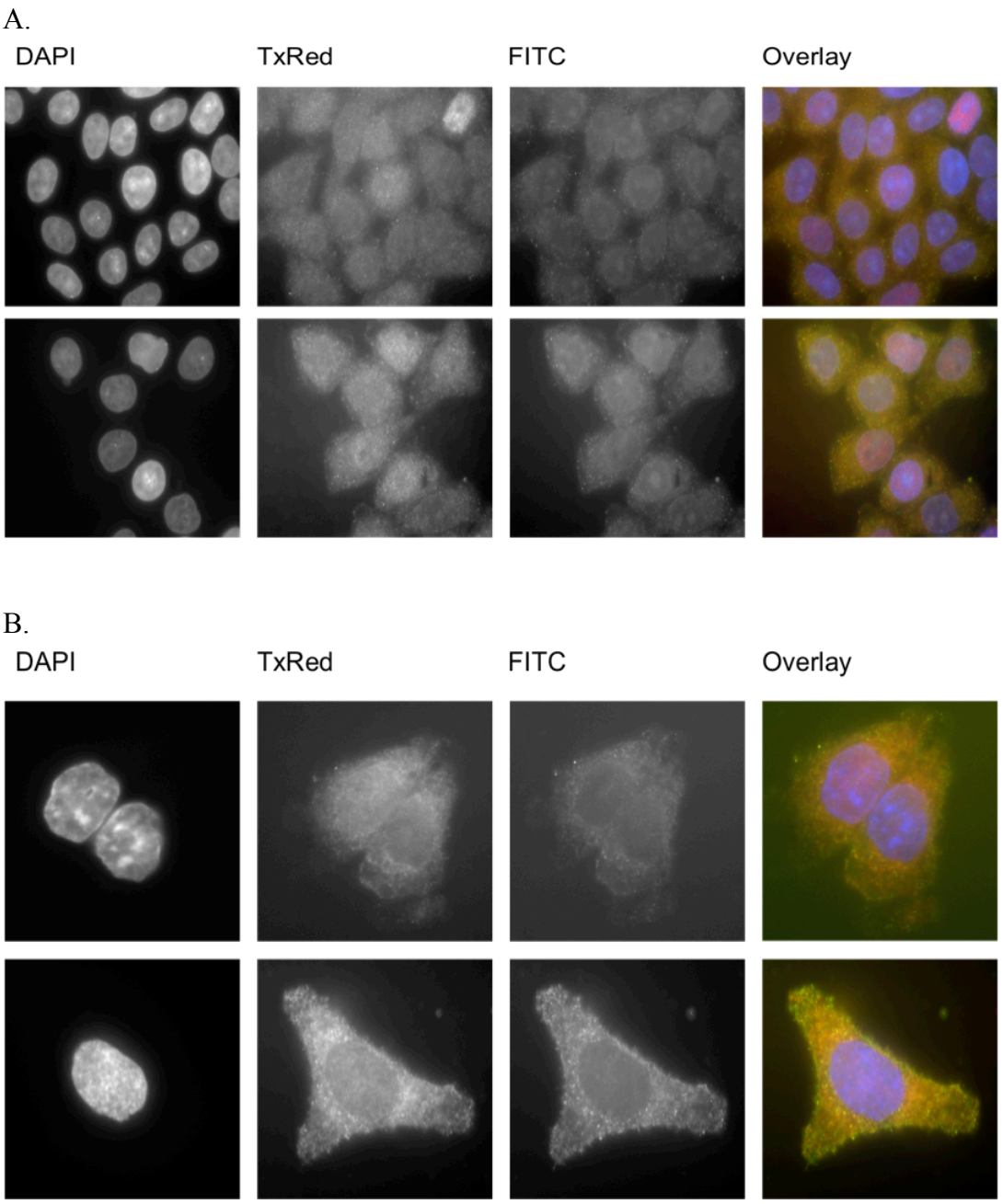


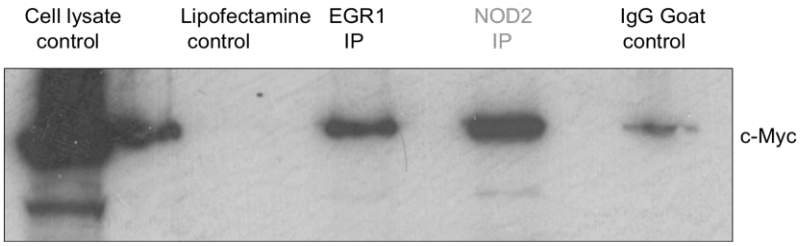
Figure 4.9: Co-localisation of EGR1 and NOD2 in CRC cell lines

HRT18 **(A)** and SW480 **(B)** cells were grown on cover slips and fixed with formaldehyde. The cells were immunostained with NOD2 anti-mouse and FITC-conjugated Mouse Anti-Sheep IgG - Horseradish Peroxidase and EGR1 anti-goat and TxRed-conjugated Goat Anti-Rabbit IgG - Horseradish Peroxidase, with DAPI to stain the nuclei (A. magification x40, B. magnification x100).

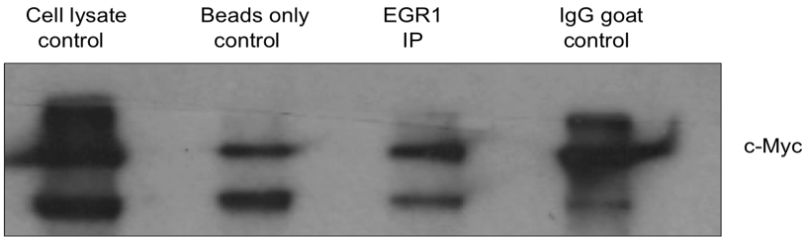
4.3.3.3 Optimisation of Immunoprecipitation of EGR1 and NOD2

A potential interaction had been identified between EGR1 and NOD2 by yeast two-hybrid analysis, hence we needed to confirm that this interaction was real and therefore developed an immunoprecipitation protocol to assess the interaction. Initially the experiments were conducted using a c-Myc-tagged NOD2 construct that was transfected into SW480 cells using Lipofectamine for 24 hours and immunoprecipitated using anti-EGR1 and agarose G beads (Figure 4.10). The membranes were then probed with anti-Myc to determine if any interaction occurs. Although there does appear to be a band in the EGR1 IP at the correct size for c-Myc, this band is also visible to a lesser extent in the IgG goat IP, which should act as a negative control for the anti- EGR1 goat antibody (Figure 4.10 A). Due to the presence of these non-specific bands in the IgG negative control, a range of experimental conditions were investigated. A control sample containing only agarose G beads was used, which showed a high level of background in the bead only sample (Figure 4.10 B). To eliminate this non-specific binding, the agarose G beads were blocked in a buffer containing BSA. (Figure 4.10 C) This resulted in a reduction of the non-specific background. The experiments were also conducted using agarose A beads which showed no difference in the levels of non-specific backgrounds.

A.



B.



C.

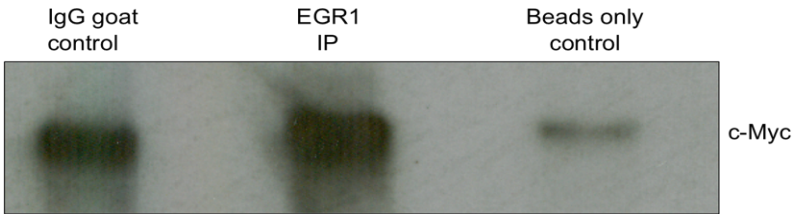


Figure 4.10: Optimisation of immunoprecipitation

SW480 cells were transfected with a c-Myc tagged NOD2 construct using Lipofectamine. Cells were extracted using lysis buffer 1 and immunoprecipitated using agarose G beads and anti-EGR1. IgG goat was used as a negative control. **A.** The immunoprecipitated complexes were run on a 10% resolving gel and probed with anti-c-Myc. **B.** A sample containing beads only was included as a negative control. **C.** The agarose G beads were cleared in a lysis buffer containing BSA to eliminate non-specific binding.

Although the results in general indicated that there was indeed an interaction between EGR1 and NOD2, it was apparent however that using the c-Myc-tagged NOD2 construct was resulting in a high level of background in the IgG control. As such it was decided to continue the experiments looking at the endogenous NOD2 protein. Due to the poor quality of the NOD2 antibodies available, the immunoprecipitation was conducted by pulling down with NOD2 and probing the membrane with anti-EGR1. A series of different buffers was investigated using varying concentrations of different salts (NaCl and KCl) and different detergents (NP40, Triton x100 and Tween) in order to eliminate any non-specific binding while retaining the interaction (Section 2.6, table 2). Two of the buffers, buffer 2 and 4, showed greatly reduced background in the IgG control with a faint band being detected as Figure 8.A illustrates for buffer 4. In order to increase the protein levels of both EGR1 and NOD2, the cells were treated with 1μg of LPS for 24 hours before extraction. As Figure 8.B illustrates the treatment with LPS does increase the endogenous EGR1 levels, and there is a faint interaction detected using buffer 2. There is also an interaction detected with buffer 4, although there is once again background in the IgG mouse control. However it was decided to continue the experiments using buffer 4, coupled with a more rigorous washing step following formation of the immuno-complex. Buffer 4 contains 300mM NaCl, with 1% Triton x100 and 50mM Hepes buffer (pH 7.4) with the addition of protease inhibitors. As Figure 8.C illustrates there is an interaction detected between EGR1 and NOD2 in both untreated and LPS treated cells.

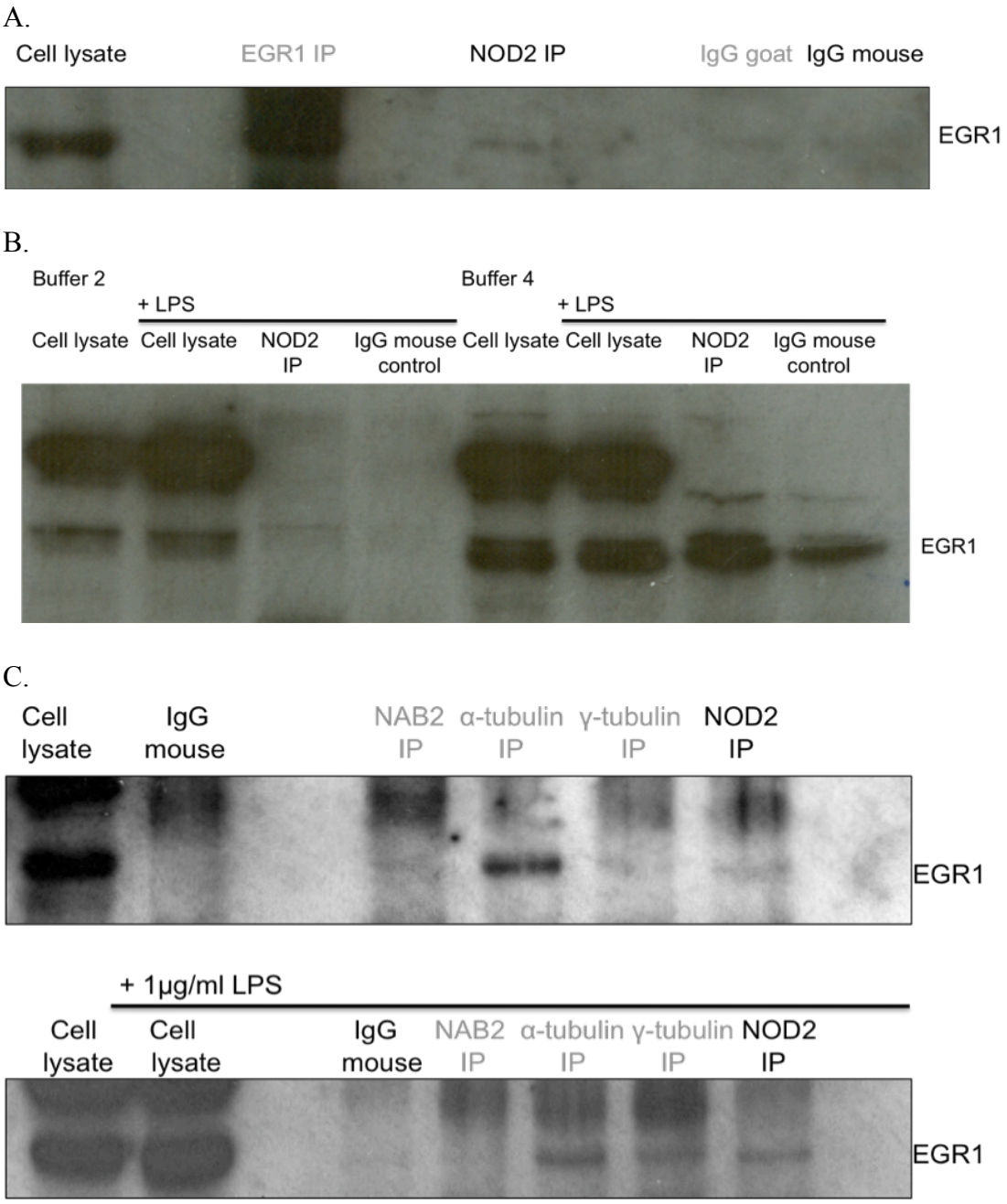


Figure 4.11: EGR1 immunoprecipitates with NOD2

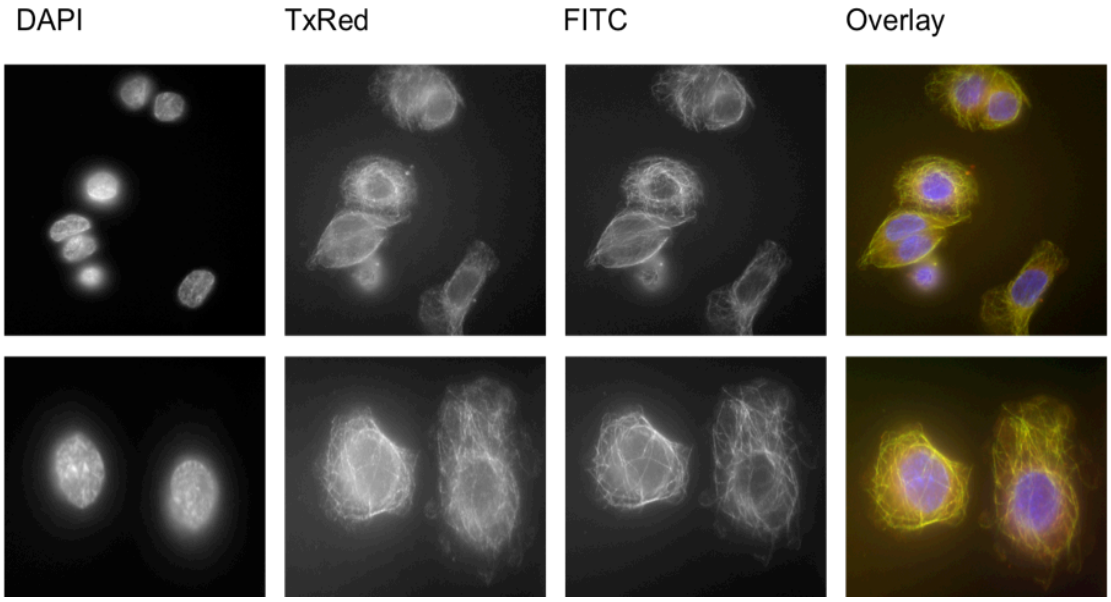
A. Cells were extracted using lysis buffer 4 and immunoprecipitated using agarose G beads and anti-EGR1. IgG goat was used as a negative control. Immunoprecipitated complexes were run on a 10% resolving gel and probed with anti-EGR1. The agarose G beads were blocked in a lysis buffer containing BSA to eliminate non-specific binding. **B.** Cells were treated with 1µg LPS for 24 hours, extracted using lysis buffer 2 and 4 and immunoprecipitated as above. **C.** Cells were extracted as above using buffer 4 and pulled down with several antibodies.

4.3.4 Investigation into EGR1 interaction with the cytoskeleton

4.3.4.1 EGR1 co-localises with tubulin in CRC cells

EGR1 has been shown to localise with microtubules, with both the α- and β-tubulin, in benign prostate cells, whereas no interaction is visible in malignant prostate cells (Mora, Olivier et al. 2004). To determine if EGR1 localises with components of the cytoskeleton in CRC cells, immunofluorescence experiments were performed. SW480 cells were co-stained with tubulin (α- and γ-tubulin) and EGR1. As can be see in Figure 4.12 EGR1 co-localises with both α- and γ-tubulin in SW480 cells. In the previous stainings of EGR1 the staining appears similar both when stained only with EGR1 and while co-staining with EGR1 and NOD2 and to some extent with the co-staining of EGR1 and γ-tubulin. However in Figure 4.12 A below the pattern of EGR1 staining when co-stained with α-tubulin appears to be quite different, both from the previous staining as seen in this chapter and from the EGR1/α-tubulin seen in the Mora, et al (Mora, Olivier et al, 2004). The difference in staining of EGR1 between these studies and ones conducted by Mora et al may be in part accounted for by the use of a different cell line as well as the use of the different antibody (Santa-Cruz sc110). However it is also likely that there is some "bleed through" of the α-tubulin antibody into the TxRed channel, giving rise to the slightly different pattern of EGR1 staining as seen in figure A below.

A.



B.

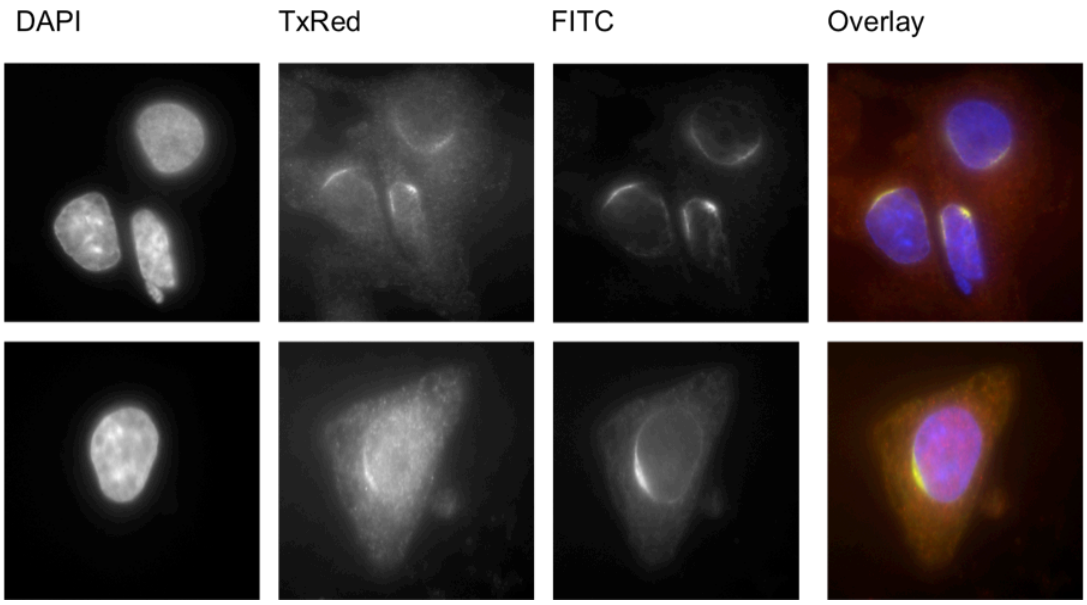


Figure 4.12: EGR1 co-localises with tubulin

A. SW480 cells were grown on cover slips and fixed with formaldehyde. The cells were immunostained with α-tubulin anti-mouse and FITC-conjugated Mouse Anti-Sheep IgG - Horseradish Peroxidase and EGR1 anti-goat and TxRed-conjugated Goat Anti-Rabbit IgG - Horseradish Peroxidase , with DAPI to stain the nuclei. **B.** Cells were grown as above and immunostained with γ-tubulin anti-mouse and FITC-conjugated Mouse Anti-Sheep IgG - Horseradish Peroxidase and EGR1 anti-goat and TxRed-conjugated Goat Anti-Rabbit IgG - Horseradish Peroxidase (magnification x100).

4.3.4.2 EGR1 can form a complex with microtubule components

In order to confirm a localisation/interaction with EGR1 and the microtubules components, α- and γ-tubulin, immunoprecipitation experiments were performed using the conditions that were optimised in the previous section (4.3.3.3). SW480 cells were treated with LPS for 24 hours before extracting with lysis buffer 4. The cellular protein extracts were precleared in BSA blocked beads, and pulled down with α- or γ-tubulin overnight. The immuno-complex was formed by adding Agarose G beads (pre-blocked in BSA). The samples were run on a 10% resolving gel and probed with anti-EGR1. As can be seen in figure 4.13 A. there is a clear interaction with EGR1 and α-tubulin in the treated cells, and a faint interaction evident with γ-tubulin. However by stimulating and inducing EGR1 levels with LPS treatment there is a clear interaction with both α- and γ-tubulin (Figure 4.13 B).

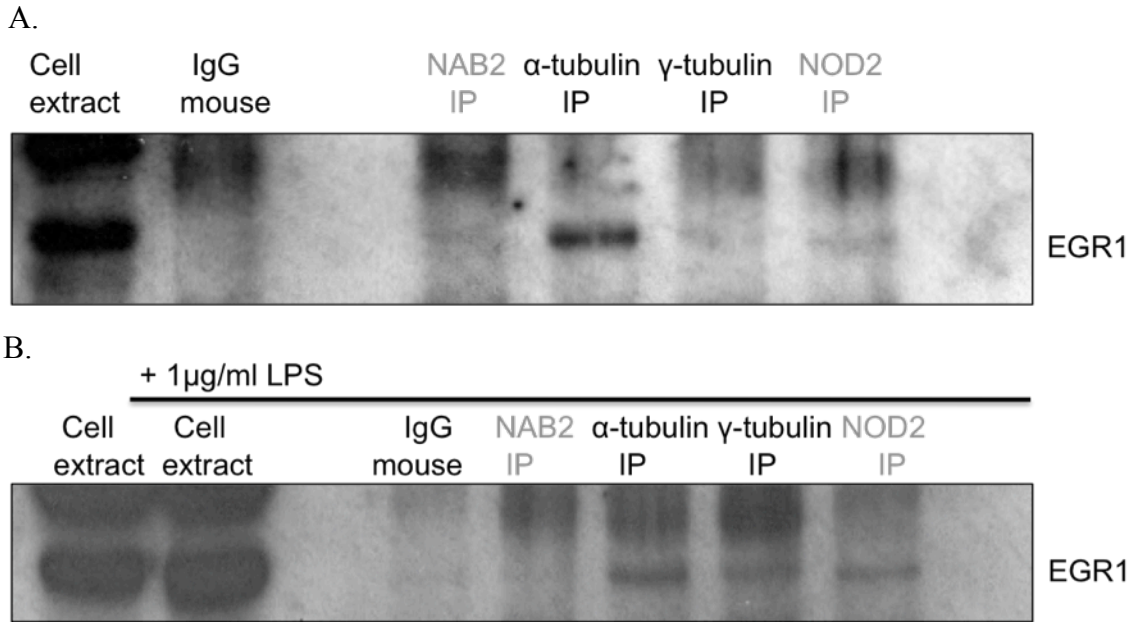


Figure 4.13: EGR1 immunoprecipitates with α- and γ-tubulin

SW480 cells were treated with 1µg LPS for 24 hours, extracted using lysis buffer 4 and immunoprecipitated using agarose G beads and anti-α-tubulin and anti-γ-tubulin. IgG mouse was used as a negative control. Immunoprecipitated complexes were run on a 10% resolving and probed with anti-EGR1. The agarose G beads were blocked in a lysis buffer containing BSA to eliminate non-specific binding. **A.** Immunopreciptation with untreated SW480 cells. **B.** Immunoprecipitation with LPS treated SW480 cells.

4.4 Discussion

The results in this chapter show that there is a high level of EGR1 protein in SW480 and HCT116 colon cancer cell lines, with lower levels in HRT18 and Vaco425 colon cancer cell lines. For the most part, this is in agreement with the qRT-PCR expression data in the previous chapter, where there are higher relative EGR1 mRNA expression levels in SW480 than HRT18. However Vaco425 cells had the highest relative EGR1 mRNA expression levels, but there appears to be less EGR1 protein expressed although it does appear to be present evenly in the cytoplasm and the nucleus of Vaco425 cells. SW480 was chosen as the main cell line in which to conduct the majority of the experiments in this chapter due to the fact that it was a SW480 cell line library that was used in the original yeast-two hybrid screen that pulled out an interaction with EGR1 and NOD2. Ideally the subcellular localisation and IP experiments would be repeated using other cell lines, in particular HRT18, which has strong levels of cytoplasmic EGR1, and Vaco425, which shows high EGR1 mRNA expression levels but lower protein levels. Due to technical difficulties and time constraits it was not possible to conduct the experiments using these cell lines but it would be important to ultimately conduct these experiments after further optimisation of the techniques. It may be possible to further stimulate expression of EGR1 and NOD2 in these cells lines, which may allow for better detection of the protein in order to look at the subcellular localisation and conduct IP experiments.

As expected EGR1 is localised predominantly in the nucleus in the CRC cells lines however there is strong cytoplasm staining evident, especially in HRT18 cells. This is in agreement with published data that demonstrates that EGR1 can localise in the cytoplasm also. EGR1 contains a bipartite nuclear localisation signal, that is located in the zinc-finger domain of EGR1 (Gashler, Swaminathan et al. 1993), and it has been suggested that the cytoskeleton may have a role to play in the nuclear translocation of EGR1 in prostate cells (Mora, Olivier et al. 2004). Mora et al, 2004 showed that EGR1 interacts with microtubules in benign prostate cells only and this interaction is necessary for the nuclear translocation of EGR1 and subsequent transcription of EGR1-mediated genes. In SW480 cells there does appear to be

higher levels of EGR1 in the cytoplasm at some time points suggesting that EGR1 may actively translocate between the cytoplasm and nucleus in CRC cells also. It is therefore of interest that we demonstrated an interaction between EGR1 and α- and γ-tubulin in SW480 cells. The interaction between EGR1 and tubulin is observed using high salt buffer conditions, which is in contrast to the published data describing an interaction with EGR1 and tubulin, where low salt with no detergent was used in prostate cells (Mora, Olivier et al. 2004). Experiments were conducted to determine if EGR1 also localises with β-tubulin, however we were unable to demonstrate an interaction. As EGR1 has been previously demonstrated to interact with β-tubulin (Mora, Olivier et al. 2004), it may be that further optimisation or an alternative β-antibody is required, or that this particular interaction does not occur in CRC cells.

As mentioned, the staining of EGR1 in some experiments is not consistent, and therefore should be interpreted with caution. For the most part the pattern of EGR1 expression in the immunflorescence is consistent and similar. However, there are some differences between experiments, particularly when EGR1 is co-stained with α-tubulin. The staining of EGR1 here is not consistent with the EGR1 staining in the study by Mora et al. However, as mentioned, they have conducted their experiments using prostate cancer cells, PC-3, LNCaP and DU-145. It is likely that there is some "bleed through" of the α-tubulin staining into the EGR1 staining. Both of the secondary antibodies were test alone to ensure that they did not show any sign of fluorescnec in the absence of a primary antibody, but it may be that the α-tubulin staining of the cells is too strong and the antibody needs to be tested using a lower concentration. One significant difference between the two studies is the EGR1 antibody. The Mora et al study used the Santa Cruz antibody Sc-110 for their immunoflorescence. In this study the same antibody (EGR1, R&D Systems AF2818) was used for both the western blotting and the immunofloorescence, and although the sc-110 EGR1 antibody did not produce good results using a Western Blot method (Figure 4.1) it could be that this antibody does not work well to detect denatured antibody and may be more successful at detecting native EGR1.

Further experiments would be needed to indicate whether the cytoskeleton actively translocates EGR1 to the nucleus in CRC cells or if there is aberrant translocation of EGR1 to the nucleus in CRC cells similar to prostate cells. Indeed it has been shown that aberrant abnormal centrosomes form in HCT116 cells (Kuriyama, Bettencourt-Dias et al. 2009). It would be interesting to disrupt the microtubules using a microtubule polymerisation agent such as nocodazole, which has been shown to prevent the translocation of EGR1 to the nucleus in prostate cells (Mora, Olivier et al. 2004), to determine if this has an effect on the localisation of EGR1 or the transcriptional activity and activation of EGR1 downstream targets in colorectal cancer cells. There are several luciferase reporter plasmids containing the promoter of genes with an EGR1 transcription factor binding site, including p300, TF and EGFR, available as a resource in the lab which could be utilised to study the effect of microtubule disruption on the transcriptional effect of EGR1. It would also be interesting to determine if there is an association with EGR1 localisation and cell cycle in CRC cells as it is known that EGR1 is involved in the regulation of genes involved in cell cycle and cell proliferation including cyclin D and thymidine kinase (Molnar, Crozat et al. 1994; Yan, Nakagawa et al. 1997).

The CD susceptibility gene/protein NOD2 is localised in the cytoplasm with some localisation apparent in the membrane bound/organelle fraction, which is in agreement with published data that suggests that NOD2 can be targeted to the cell membrane in epithelial cells after MDP stimulation, and this membrane association is necessary for NF-κB activation (Barnich, Aguirre et al. 2005). There have been several mutations of NOD2 reported (Leung, Hong et al. 2007; Rosenstiel, Huse et al. 2007; McCarroll, Huett et al. 2008), the most common being a truncated protein caused by a frame shift mutation, Leu1007fins, which is associated with Crohn's disease and does not respond to MDP or activate NF-κB (Ogura, Bonen et al. 2001). This CD-associated mutation has an impaired ability to activate NF-κB, and fails to associate with the cell surface membrane. The two other commonly CD-associated polymorphisms, Arg702Trp and Gly908Arg, also result in decreased activation of NF-κB.

As discussed in the previous chapter there is significantly less EGR1 transcript present in CD patients compared with healthy controls, and a majority of patients show a reduction in EGR1 expression after treatment with LPS (70%), MDP (65%) and TNF (66%), suggesting that there is aberrant inflammatory response in these patients. It is interesting then that we have demonstrated an interaction between EGR1 and NOD2 in SW480 cells. The interaction between EGR1 and NOD2 is observed using high salt buffer conditions, in both untreated and LPS treated cells. It should be noted that the LPS used in the immunoprecipitation experiment is known to be contaminated with MDP, unlike the LPS that was used for the qRT-PCR expression analysis and therefore could be MDP stimulation that was influencing the EGR1/NOD2 interaction. It is probable that the interaction between EGR1 and NOD2 occurs in the cytoplasm, however there is evidence to suggest that NOD2 is membrane associated and as EGR1 is also present in the membrane bound/organelle fraction there is the possibility that the interaction is membrane associated also. To further investigate this it would be interesting to determine if EGR1 interacts with any of the three CD-associated NOD2 polymorphisms. HA-tagged constructs of the three common NOD2 mutations are available as a resource (gift from Colin Stevens). Preliminary experiments conducted to transfect these constructs into CRC cells to determine an interaction with EGR1 by immunoprecipitation were unsuccessful, however if conditions were optimised to reduce the non-specific binding that occurred using transfected constructs, it would be interesting to determine if EGR1 co-localises and/or interacts with any of the NOD2 mutants. This would also provide more information as to whether or not the LRR domain of NOD2 is required for the EGR1-NOD2 and it would also be interesting to identify the physical regions that are necessary for the EGR1-NOD2 interaction. Construsts lacking domains or regions of either protein could be used to determine which domains are necessary for the interaction.

EGR1-protein interactions can be complicated and it has been shown that the stimuli under which EGR1 is induced can have different effect on its post-translational modification, indeed sometimes with opposing effects. For example in prostate cells, serum-induced EGR1 can be acetylated which then can bind to p300 /CBP and

results in the inhibition of p300/CBP and EGR1 itself (Yu, de Belle et al. 2004; Adamson, Yu et al. 2005). However if EGR1 is induced by UV irradiation, the protein become phosphorylated which then transactivates p53, p73 and PTEN and inhibits NF-kB. It would appear that acetylated EGR1 is growth promoting and anti-apoptotic whereas the opposite is true for phosphorylated EGR1. It would be very interesting to determine if either acetylation or phosphorylation of EGR1 is necessary for EGR1-NOD2 interaction. It is also unknown as to what effect the EGR1-NOD2 interaction may have on either EGR1 or NOD activation of downstream targets, whether stimulatory, inhibitory or indeed if it has any effect on the activation of the EGR1 or NOD2 dependent pathways. An interesting experiment would be to determine what effect loss of NOD2 would have on the ability of EGR1 to activate gene transcription using the EGR1 luciferase reporter plasmids with EGR1 binding sites or whether loss of EGR1 effects the MDP-mediated NOD2 dependent activation of a NF-κB reporter assay. This could lead to a greater understanding of the role of both EGR1 and NOD2 in inflammatory pathways that are known to be important in both IBD and CRC.

5 Chapter 5

5.1 Introduction

Little has been reported on the protein-protein interactions of EGR1, particularly in epithelial cells and CRC cells. However EGR1 protein interactions are known to have important cellular functions. EGR1 interacts with its inhibitor proteins, NAB1/NAB2, which helps regulate EGR1 expression levels in cells (Swirnoff, Apel et al. 1998). As previously discussed (Ch. 4), EGR1 is also known to interact with components of the cytoskeleton in prostate cancer which has been suggested to play a role in EGR1 translocation (Mora, Olivier et al. 2004). Finally acetylated EGR1 is known interact with p300/CBP, also in prostate cells, and inhibit its function (Adamson, Yu et al. 2005).

We have demonstrated that EGR1 is differentially expressed in colonic epithelial cells, in both CRC and IBD patients, as well as confirming a novel EGR1-NOD2 interaction in SW480 cells, which was first identified by yeast-two hybrid screen. Therefore it is of interest to investigate other potential EGR1-protein interactions in epithelial cells in order to identify novel interactions and new pathways which may be involved in colorectal disease to further our understanding of disease.

It was decided to investigate whether potential EGR1 protein interactions could be indentified using a yeast-two hybrid screen, using EGR1 as bait and with a SW480 cDNA library in order to determine any novel EGR1 interacting proteins in cancer epithelial cells. The screen identified 22 potential novel interactions with EGR1. Two interesting proteins identified by the screen, NPM1 and eEF1A1 were further investigated. The localisation and expression of both proteins was determined in CRC cells and the interactions confirmed by immunoprecipitation.

5.2 Methods

The yeast-two hybrid screen was conducted using Clontech Matchmaker^{TM} system.

5.2.1 Background of a yeast two-hybrid screen

The yeast two hybrid screen is an assay to detect protein-protein interactions. It works by utilising the binding of the GAL4 transcription factor to an upstream activating sequence (UAS) activating the transcription of downstream reporter genes such as lacZ. The GAL4 transcription factor consists of a DNA binding domain (BD) which binds to the UAS and a activating domain (AD) which is responsible for the activation of the transcription factor (Figure 5.1 A).

A yeast two hybrid screen takes advantage of this system by fusing a 'bait' protein of interest, in this case EGR1, to the GAL4 DNA-binding domain, and a cDNA library 'prey' is fused to the GAL4 activation domain. Expression of either the bait+DNA BD or the prey+AD in yeast cells is inadequate to activate gene transcription (Figure 5.1 B,C). Upon interaction of the bait protein with a prey protein, this brings the BD and AD domains in proximity, allowing activation of transcription (Figure 5.1 D).

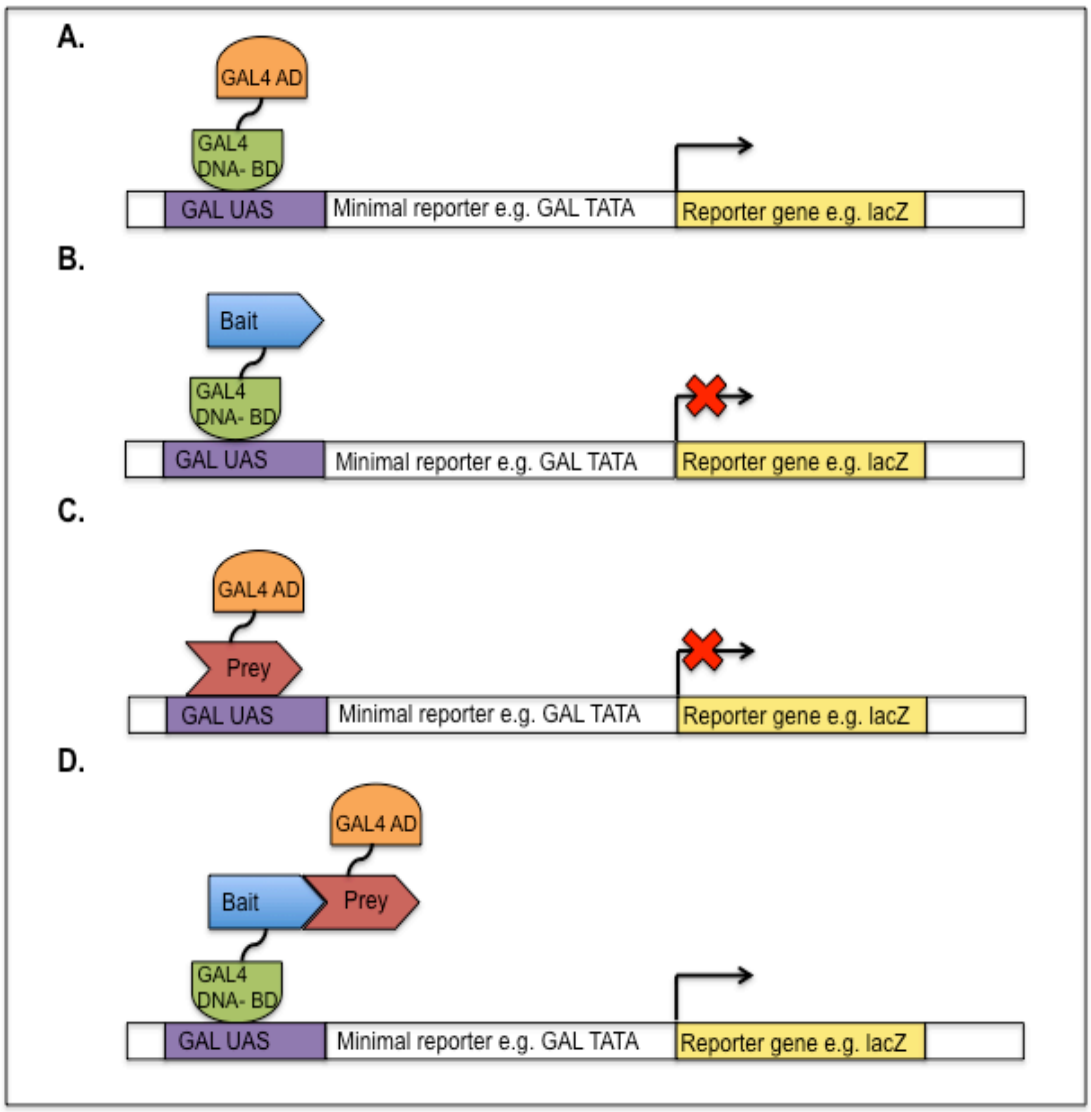
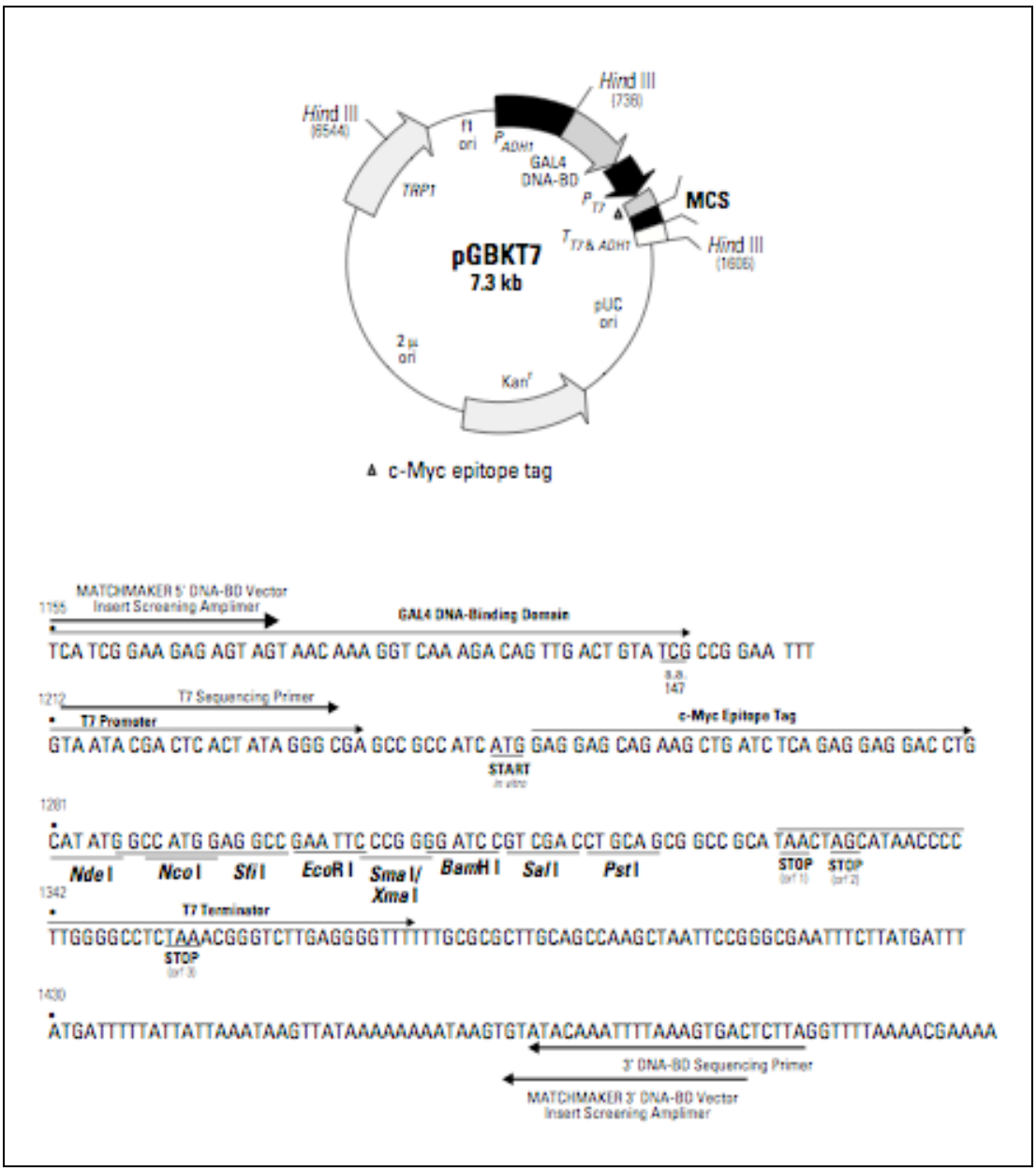


Figure 5.1: Illustration of GAL4 mediated activation of reporter genes in yeast two-hybrid system

Diagram of the activation of reporter genes by the GAL4 transcription factor in yeast two hybrid. **A**. GAL4 DNA binding domain (BD) and activation domain (AD) are both required for activation of reporter gene transcription. **B.** 'Bait' protein is fused to the DNA-BD, but cannot activate reporter genes. **C.** 'Prey' protein fused with the AD domain and also cannot activate gene transcription. **D.** If a protein-protein interaction occurs with the bait and prey proteins, the BD and AD domains can then activate transcription of the reporter genes

The bait (EGR1) was cloned into the vector pGBKT7 and transformed into the Y187 yeast strain (Figure 5.2 A). The Y187 yeast strain has a GAL1 UAS, a GAL1 TATA box and a lacZ reporter gene. The pGBKT7 vector contains the GAL4 DNA BD, a kanamycin resistance gene and the nutritional reporter TRP1, indicating that it is selected for in the absence of the amino acid tryptophan. The cDNA library (SW480 – prepared by GI unit) was cloned into the vector pGADT7 which has the GAL4 AD, an ampicillin resistance gene and the nutritional reporter LEU2, for the amino acid leucine, using the yeast stain AH109 (Figure 5.2 B). AH109 has three reporter genes, ADE2, HIS3 and MEL1 (lazZ), in order to reduce the number of false positives detected. The nutritional reporters are used to select for the approprioate vectors, the 'bait' protein in Y187 grows in the absence of tryptophan (-Trp), the 'prey' protein in AH109 grows in the absence of leucine (-Leu). Once the two yeast are mated, the presence of an interaction will active the lacZ reporter as well as the other two nutritional reporters, so will only grow in the absence histidine (-His), with more stringent interactions growing in the absence of adenine (-Ade).

A.


Hind III
(736)
Hind III
(5544)
f1
ori
P_{ADH1}
GAL4
DNA-BD
TRP1
P_{T7}
MCS
pGBKT7
7.3 kb
$T_{T7\&\ ADH1}$
Hind III
(1606)
pUC
ori
2 μ
ori
Kan^r
▲ c-Myc epitope tag
MATCHMAKER 5' DNA-BD Vector
Insert Screening Amplimer
1155
GAL4 DNA-Binding Domain
TCA TCG GAA GAG AGT AGT AAC AAA GGT CAA AGA CAG TTG ACT GTA TCG CCG GAA TTT
aa
147
1212
T7 Sequencing Primer
T7 Promoter
c-Myc Epitope Tag
GTA ATA CGA CTC ACT ATA GGG CGA GCC GCC ATC ATG GAG GAG CAG AAG CTG ATC TCA GAG GAG GAC CTG
START
in vitro
1281
CAT ATG GCC ATG GAG GCC GAA TTC CCG GGG ATC CGT CGA CCT GCA GCG GCC GCA TAACTAGCATAACCCC
Nde I
Nco I
Sfi I
EcoR I
Sma I/
Xma I
BamH I
Sal I
Pst I
STOP
STOP
1342
T7 Terminator
TTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCGCGCTTGCAGCCAAGCTAATTCCGGGCGAATTTCTTATGATTT
STOP
1400
ATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAA
3' DNA-BD Sequencing Primer
MATCHMAKER 3' DNA-BD Vector
Insert Screening Amplimer

B.

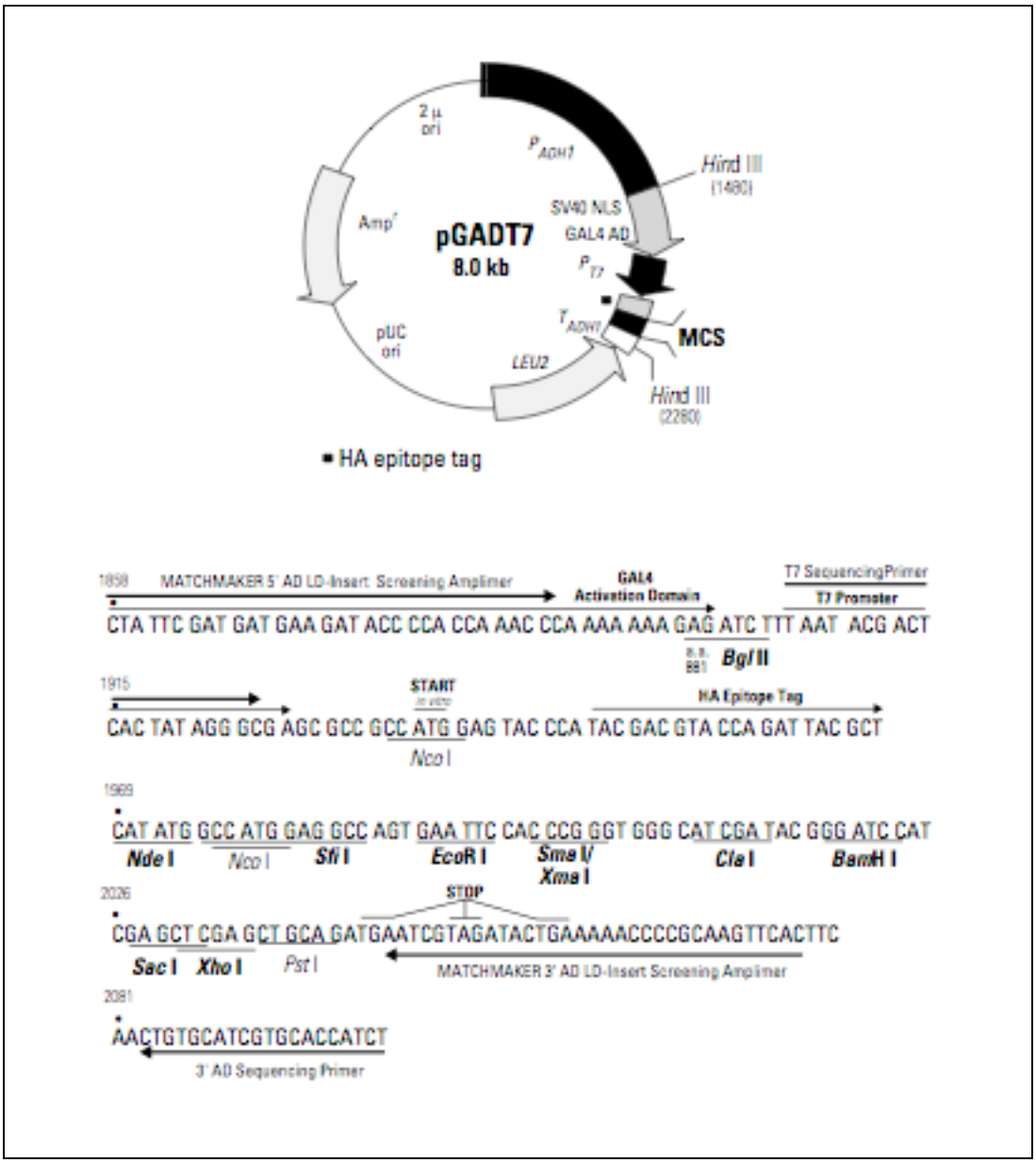


Figure 5.2: Map of the vectors pGBKT7 and pGADT7

Restriction map and multiple cloning sites of pGBKT7 (A; for the 'bait') and pGADT7 (B; for the 'prey') vectors that are use in the yeast two hybrid screen. Vector maps are from Clontech Laboratories Inc (Clontech Laboratories Inc 2008; Clontech Laboratories Inc 2009).

5.2.2 Cloning of *EGR1* into yeast vector pGBKT7

5.2.2.1 Cloning of *EGR1* into pGEM-T cloning vector

As transcription factors can activate transcrisption of the reporters in the yeast without the presence of an interacting protein, an *EGR1* construct without its transactivation domain as well as a full length *EGR1* construct were cloned into the pGBKT7 vector. Full length *EGR1* cDNA (EGR1Fl) and *EGR1* minus its transactivation domain (EGR1ΔAct) were amplified from CRC cell line cDNA and the PCR products were purified using a PCR Purification Kit (Qiagen). EGR1Fl and EGR1ΔAct PCR products were cloned into the pGEM-T cloning vector (Promega), and transformed into 50μl of TOP10 Chemically Competent *E.coli* cell (Invitrogen). Colonies were selected and amplified using Sp6 and T7 primers (Ch.2.4.1). A colony that was successfully amplified, producing a single band, was grown in 250ml L-Broth, with 20mg/ml ampicillin, overnight with shaking at 37^{o}C. Both plasmid inserts were extracted, sequenced to confirm correct construct and purified using a Maxi Prep Kit (Qiagen) as per manufacturer's instructions.

5.2.2.2 Cloning of *EGR1* into pGBKT7

The constructs containing EGR1Fl and EGR1ΔAct in pGEM-T and the pGBKT7 vector were cut using *EcoR1* in a 100μl reaction at 37^{o}C for 2 hours. Calf Intestine Alkaline Phosphatase (CIAP) (Roche) was added to the reactions to prevent self re-ligation. The restriction digest reactions were run out on an agarose gel (0.8-1.5%), with the correct band extracted and the DNA purified using a Gel Extraction Kit (Qiagen). A ligation reaction (20μl) was set up using 1μg of cut pGBKT7 vector and 3μ of cut insert. The reaction was incubated at room temperature for 2 hours. 5μl of this ligation reaction was transformed into 50μl of TOP10 Chemically Competent *E.coli* cells (Invitrogen) and grown overnight at 37^{o}C. Colonies were selected and amplified using primers A308 and A309 (section 2.4.1), and then sequenced to confirm the orientation of the insert. A colony with the correct sequence and orientation was grown in 250ml L-Broth, with 10mg/ml kanamycin, overnight with shaking at 37^{o}C. The EGR1FL plasmid and the EGR1ΔAct plasmid were extracted

and purified using Maxi Prep Kit (Qiagen) as per manufacturer's instructions, and accurate plasmid colonies stored as glycerol stocks at –80^{o}C, via addition of 1ml of overnight culture to 300µl sterile glycerol.

5.2.3 Yeast-two hybrid screen

The yeast strains Y187 and AH106 were grown in YPD medium or on YPD agar plates. SD media and agar with appropriate nutritional additives were used to grow the yeast strains transformed with the bait or library. All yeast was incubated at 30^{o}C and grown for 3-5 days. Details of the nutritional additives are found in chapter 2.7, with a summary below.

SD yeast media/agar	SD media/agar with 10x dropout solution + 10x histidine, 10x adenine, 10x leucine, 10x tryptophan
SD/-Trp media/plates	SD media/agar with 10x dropout solution + 10x histidine, 10x adenineand 10x leucine
SD/-Leu/-Trp media/plates	SD media/agar with 10x dropout solution + 10x histidine and 10x adenine
SD/-His/-Leu/-Trp media/plates	SD media/agar with 10x dropout solution + 10x adenine
SD/-His/-Ade/-Leu/-Trp media/plates	SD media/agar with 10x dropout solution

Table 5.1: Summary of the yeast SD media/plates with nutritional additives

5.2.3.1 Lithium Acetate (LiAc) Yeast Transformation

The EGR1Fl and EGR1ΔAct constructs in pGBKT7 were transformed into the yeast strain Y187 using a LiAc transformation. The constructs were grown overnight in YPD media at 30^{o}C with shaking. 300ml of YPD was inoculated with 30ml of this starter culture to give an Optical Density (OD at 600nm) of ~0.2-0.3. The culture was grown until it reached an OD of 0.4-0.5. The cells were harvested by centrifugation at 1000xg for 5min, resuspended in 50ml H_2O, spun again at 1000g for 5 min and resuspended in 1.5ml sterile 1 x TE/1 x LiAc solution.

0.1µg plasmid was added to 100µg salmon sperm, 100µl of Y187 yeast cells and 0.6ml sterile PEG/LiAc solution and vortexed for 10sec. The cells were incubated at 30°C for 30min, after which 70µl of DMSO was added and then placed in 42°C water bath for 15min to heat shock. The cells were chilled on ice for 2 min, centrifuged briefly at 14,000 rpm and resuspended in 0.5ml sterile 1x TE. A 1/10 dilution of the cells (100µl) was plated on SD/-Trp plates and incubated at 30°C for 3 days.

5.2.3.2 Test Mating

50ml of SD/-Trp media were inoculated with the transformed constructs and grown overnight at 30°C with shaking. The cultures were grown until they reached an OD of 0.7-1.0. The cells at a concentration of $4x10^8$/ml were mixed with 250µl of SW480 library cells in AH109 ($2x10^8$/ml) and plated out onto 2 YPDA plates and incubated overnight at 30°C. The SW480 library cells in AH109 were provided by E.Nimmo from the GI unit. 5ml of YPDA media was added to each plate and the cells were scraped off, centrifuged at 1000g for 5 min and resuspended in 2.5ml of SD/-Leu/-Trp media. The following dilutions were made and plated out onto the relevant plates (Table 5.2). The mated cells will grow on the SD/-Leu/-Trp, SD/-Leu and SD/-Trp plates, which will be used to calculate the mating efficiency, but only colonies containing interaction proteins will activate transcription of the reporter genes and grow on the plates without adenine and histidine (SD/-His/-Leu/-Trp; SD/-His/-Ade/-Leu/-Trp).

Dilution	**Amount**	**Plates:**
Undiluted	240ul	SD/-His/-Leu/-Trp; SD/-His/-Ade/-Leu/-Trp
1/10	100ul	SD/-Leu/-Trp
1/1000	100ul	SD/-Leu/-Trp; SD/-Leu; SD/-Trp
1/100000	100ul	SD/-Leu; SD/-Trp

Table 5.2: Dilutions of Y2H mating plated

The cells were grown for 3-5 days at 30^{o}C and the mating efficiency was calculated. Colonies were picked from the SD/-His/-Leu/-Trp and the SD/-His/-Ade/-Leu/-Trp plates and inoculated into 96 well plates containing YPD + 25% glycerol and frozen at –80^{o}C.

5.2.3.3 Products of Yeast Mating

Colonies resulting from the yeast mating of the strain containing the SW480 library and strain containing the EGR1Fl and EGR1ΔAct constructs were amplified by PCR and visualised on an agarose gel. PCR products which produced a single band were purified and sequenced. The DNA sequences were translated into protein sequences and put into BLAST (Basic Local Alignment Search Tool; NCBI) to identify the proteins whose sequence was amplified from the mated yeast colonies.

5.2.4 Detection of false positives

Although the use of the three reporter genes, LacZ, ADE2 and HIS3 reduce the number of false positives detected, it is still important to ensure that the 'prey' protein detected does not activate transcription of the reporter genes by itself. The interacting prey plasmids were isolated from the yeast (Chapter 2.17.5) and were co-transformed into AH109 yeast cells with the pGBKT7-pGBKT7 construct, and with a pGBKT7 construct only using a LiAc transformation as before. Transformed cells were plated onto SD/-Leu/-Trp/x-gal and SD/-Ade/-Leu/-Trp/x-gal plates to determine if the prey only constructs activate the reporter genes.

5.3 Results

5.3.1 Cloning of EGR1 constructs into the pGBKT7 vector

Two EGR1 proteins constructs were investigated in the yeast-two hybrid system, a full length EGR1 (EGR1Fl) and a truncated EGR1 protein lacking the transactivation domain (EGR1ΔAct) as it is known that the transactivation domain of transcription factors can bind to the GAL4 UAS and activate transcription of the reporter genes in yeast two hybrid thereby generating false positives. Studies indicate that the transactivation domain of EGR1 consists of amino acid 1-~285, followed by an inhibitory domain (Figure 5.3) (Russo, Sevetson et al. 1995).

Primers were designed to amplify full length *EGR1* and an *EGR1* construct from amino acid codon 283, incorporating an EcoR1 restriction enzyme site. EGR1Fl and EGR1Act were amplified using HRT18 cDNA and sequenced to determine accurate amplification. Both constructs were cloned into the pGEM-T cloning vector, and then cloned into the yeast-two hybrid vector pGBKT7.

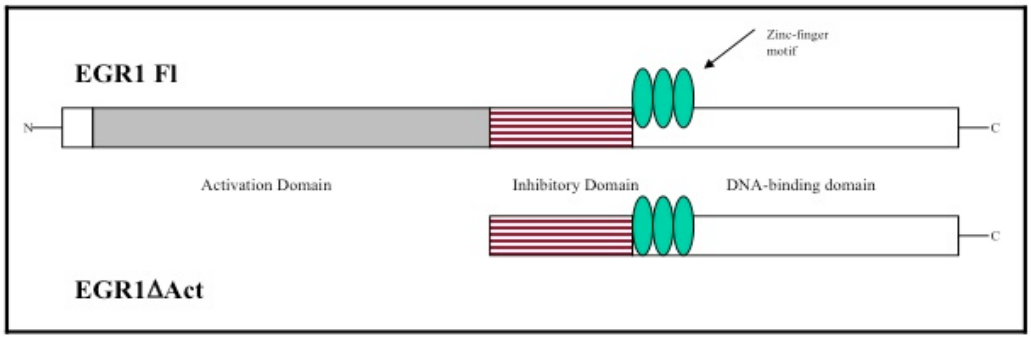


Figure 5.3: EGR1Fl and EGR1ΔAct constructs

Illustration of the EGR1 full length construct (EGR1Fl) and the EGR1 construct minus the transactivation domain (EGR1ΔAct) (Thiel and Cibelli 2002).

5.3.2 LiAc Yeast Transformation

The pGBKT7-EGR1Fl and the pGBKT7-EGR1ΔAct construct were transformed into the yeast strain Y187 using a LiAc transformation and plated on SD/-Trp agar plates, as the pGBKT7 vector has a selective for tryptophan (Trp). Two different clones for each construct were transformed, as well an empty vector. The transformation efficiencies of the constructs into the yeast strain were calculated using the formula below, where cfu is the number of colonies present (colonies forming unit).

Equation 5.1: Transformation efficiency

$$\frac{\text{Cfu x total suspension vol. (µl)}}{\text{Vol. plated (µl) x dilution factor x µg DNA used}} = \text{cfu/µg DNA}$$

Construct	Transformation Efficiency
pGBKT7 vector only	3.5×10^4 cfu/µg DNA
pGBKT7-EGR1Fl (1)	5.5×10^4 cfu/µg DNA
pGBKT7-EGR1Fl (2)	6×10^4 cfu/µg DNA
pGBKT7-EGR1ΔAct (1)	3.1×10^5 cfu/µg DNA
pGBKT7-EGR1ΔAct (2)	2.75×10^5 cfu/µg DNA

Table 5.3: Transformation efficiencies of pGBKT7-EGR1 constructs

5.3.3 Preliminary Yeast-two hybrid experiments

As the transformation efficiencies were within the accepted range, several control experiments were conducted prior to the yeast-two hybrid screen. In order to ensure that the transformed EGR1 constructs had no toxic effects on yeast growth, a toxicity assay was performed by measuring the absorbance of the transformed yeast culture over a period of time. The resulting measurements were used to plot a growth curve. The growth curves of the transformed constructs are similar to that of the Y187 strain alone, indicating that there are no toxic effects of the constructs.

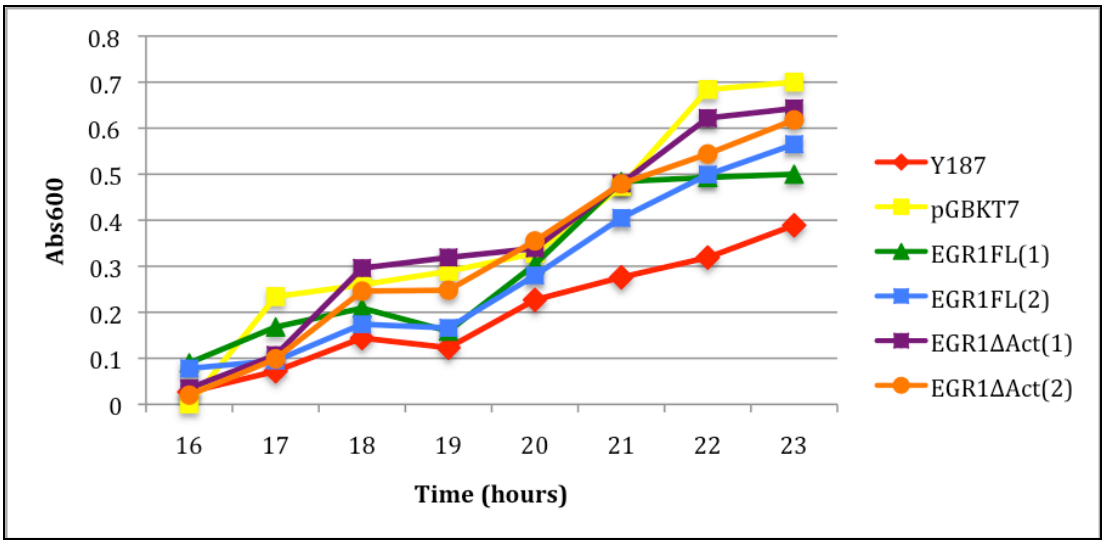
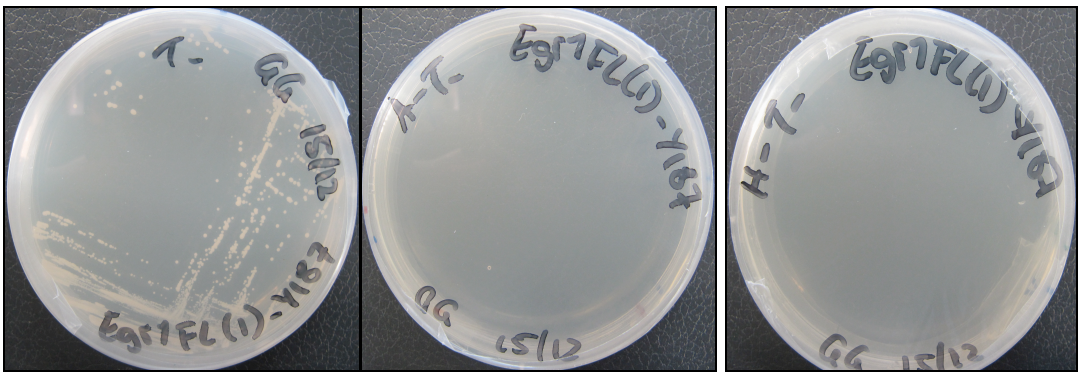


Figure 5.4: Growth curve of Y187, pGBKT7 and the EGR1 constructs

The Y187 yeast strain, the pGBKT7 vector and the transformed EGR1 constructs were grown overnight in SD/-Trp media. The absorbance at 600nm of each culture was measured every hour for 8 hours and used to plot a growth curve.

The second control experiment conducted was a self-activation test. This was performed to ensure that the transformed cells with EGR1 constructs did not activate transcription of the lacZ reporter as well as ensuring the cells only grow in the absence of tryptophan and not in the absence of histidine (His) or adenine (Ade) as the absence of these amino acids is required for determining whether the two proteins interact and the strength of this interaction. As such the transformed Y187 yeast cells did not grow on SD/-Trp/-His and SD/-Trp/-Ade plates, indicating that at this stage the transformed constructs require His and Ade to grow. Figure 5.5 demonstrates this for one of the EGR1Fl constructs but this was true for all constructs.

Y187 transformed with EGR1Fl construct



SD/-Trp plate SD/-Ade/-Trp SD/-His/-Trp

Figure 5.5: Auto-activation assay

Y187 yeast cells transformed with EGR1 constructs in pGBKT7 vector were grown on SD/-Trp/x-gal, SD/-Ade/-Trp and SD/-His/-Trp plates to determine if the bait alone activates the two reporters, ADE and HIS3. This figure illustrates the results for the EGR1Fl construct, which only grows in the absence of tryptophan and still required the presence of adenine and histidine to grow.

5.3.4 Yeast-two hybrid mating

A yeast mating was performed using one of the EGR1Fl and EGR1ΔAct constructs. The Y187 transformed constructs were mated with an AH109 yeast strain containing a SW480 cell line library (E.Nimmo). The SW480 library is transformed into the yeast vector pGADT7, which is selective for leucine (Leu). Positive interactions should activate the HIS3 reporter and the ADE2 reporter if the interaction is strong, so mated colonies were selected for growth on both SD/-His/-Leu/-Trp and SD/-His/-Ade/-Leu/-Trp. The mating efficiency was determined by inoculating the mated yeast colonies onto SD/-Leu (to determine viability of the prey library) and SD/-Trp (to determine viability of the bait) and SD/-Leu/-Trp to determine viability of the diploid. The mating efficiency was determined using the following formula:

Equation 5.2: Viable cfu/ml

$$\frac{\text{Cfu}}{\text{vol. plated (ml) x dilution factor}} = \text{cfu/ml}$$

Equation 5.3: Mating efficiency

$$\frac{\text{No. of cfu/ml of diploids}}{\text{No. of cfu/ml of limiting partner (either prey library or bait)}} \times 100 = \%\ \text{diploid}$$

The mating efficiency for the strain containing the EGR1Fl construct was 10% and 6.78% for the EGR1ΔAct construct. The interactions were confirmed by selecting the colonies on SD/-His/-Ade-/-Leu/-Trp plates. Positive colonies were scored for growth in the absence of adenine.

5.3.5 Confirmation of interactions

The presence of the lacZ reporter allows for a further test for activation of reporter genes. An X-gal assay was conducted to detect the activation of the lacZ reporter, using a colony lift filter assay. Colonies were streaked onto SD/-Leu/-Trp plates (4 per clone) and allowed to grow for 3 days a 30^{o}C. Positive and negative controls were added (provided by GI unit; Invitrogen).

LEU2 Plasmid	TRP1 Plasmid	Control
pEXP™32/Krev1	pEXP™22/RalGDS-wt	Strong positive interaction control
pEXP™32/Krev1	pEXP™22/RalGDS-m1	Weak positive interaction control
pEXP™32/Krev1	pEXP™22/RalGDS-m2	Negative interaction control
pDEST™32	pDEST™22	Negative activation control
pGADT7	pDEST™22	Negative activation control

Table 5.4: Positive and negative controls for Y2H screen (Invitrogen)

A nitrocellulose membrane was placed over the yeast colonies, removed and frozen at -80^{o}C for 1 hour. Whatman filter paper was then placed in Z buffer with X-gal solution, and the nitrocellulose was placed yeast side up on the Whatman paper to allow the solution to soak through. After an overnight incubation at 37^{o}C, the colonies were scored for blue colour. Any colonies remaining white were eliminated as false positives (Figure 5.6). The remaining blue colonies were sequenced to determine the interacting proteins. Table 5.5 lists the interacting proteins that were identified from the screen.

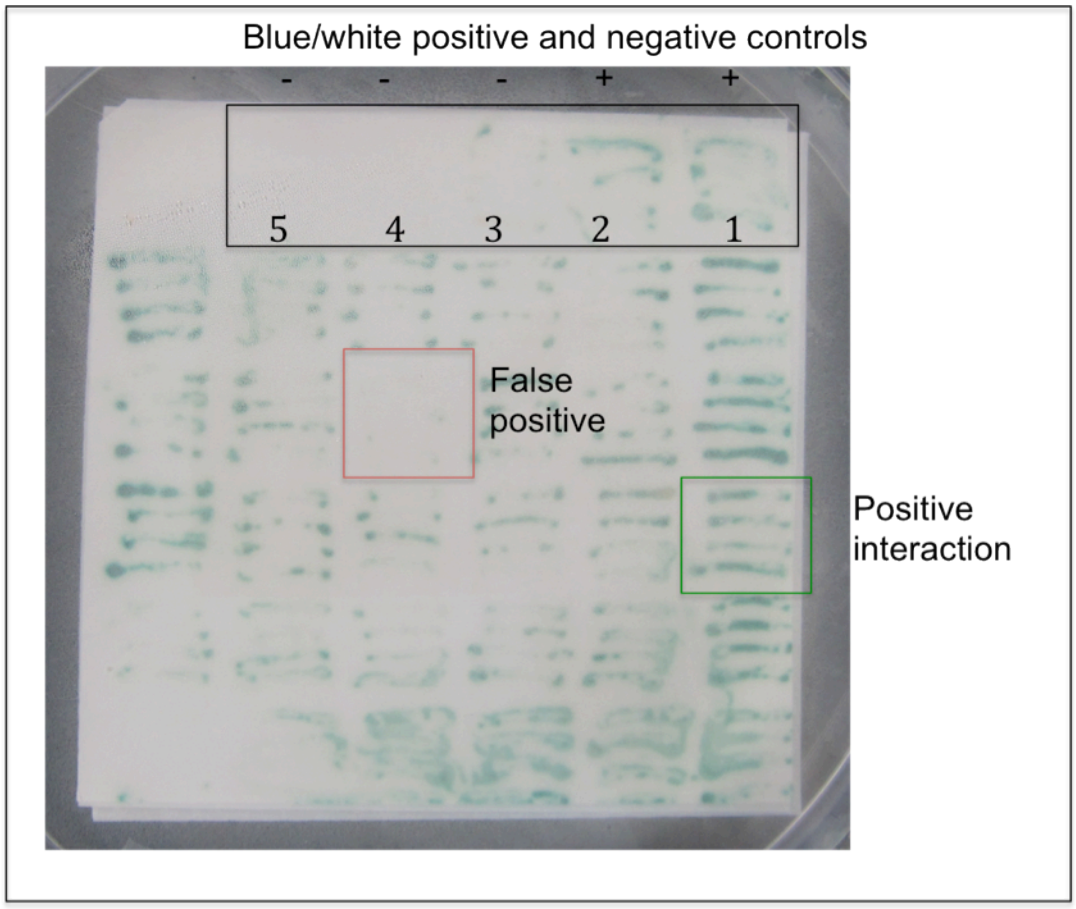


Figure 5.6: Colony lift filter assay

Yeast colonies were pressed onto nitrocellulose membrane frozen for 1 hour at -80^{o}C, and soaked in a Z buffer/X-gal solution followed by an overnight incubation at 37^{o}C to determine activation of the lacZ reporter. Yeast colonies that turn blue indicate activation of lacZ. A range of positive and negative controls were added to the top row of the membrane, 1= pEXP32/Krev1+pEXP22/RalGDS-wt; 2=pEXE32/Krev1+pEXP22/RalGDS-m1; 3=pEXP32/KREV1+pEXP22/RalGDS-m2: 4=pDEST32+pDEST22:5=pGAD7+pDEST22.

Protein Name	EGR1 Construct	No.of colonies	Strength of interaction	DAVID classification
RPL27	EGR1Fl	2	+++	Group I
RPL21	EGR1Fl	2	++	Group I
RPL5	EGR1Fl	1	++	Group I
RPL13a	EGR1Fl	2	++	Group I
RPS2	EGR1Fl	1	+++	Group I
TAF8	EGR1Fl	1	++	Group II
ZIMZ2	EGR1Fl	1	++++	Group II
NFAT5	EGR1Fl	1	+++	Group II
eEFIA1	EGR1Fl	1	+++	NA
YIF1A	EGR1Fl	2	++	NA
CLECD2	EGR1Fl	1	++	NA
PTMS	EGR1Fl	1	++	NA
NACA	EGR1Fl	1	+++	NA
PSPH	EGR1Fl	1	+++	NA
BASP1	EGR1Fl	1	+++	NA
NPM1	EGR1Fl	3	++	NA
eIF4A1	EGR1Fl	2	++	NA
FKBP	EGR1Fl	2	+++	NA
GNAS	EGR1Fl	1	+++	NA
SLC7A8	EGR1Fl	2	++	NA
PHC1	EGR1Fl EGR1ΔAct	7	+++	NA
C11orf17	EGR1Fl EGR1ΔAct	14	+++	NA

Table 5.5: Potential EGR1 interaction proteins as determined by Y2H screen.

5.3.6 Potential EGR1 interacting proteins

The yeast-two hybrid screen identified 22 novel proteins that potentially interact with EGR1. All of the proteins identified were found to interact with the full length EGR1 protein. Two proteins, PHC1 and C11orf17, interacted with both the full length and ΔAct construct. Ten of the proteins were isolated in two or more clones. The list of candidate genes was submitted to The Database for Annotation, Visualization and Integrated Discovery (DAVID), an online resource that performs a "Gene Functional Classification", separating genes into functionally related groups (Huang da, Sherman et al. 2009; Huang da, Sherman et al. 2009). Upon submission of the list of genes for the yeast-hybrid screen, two functionally similar groups were identified Five genes were identified in Group I, and three genes were identified in Group II. None of the remaining proteins were found to be functionally similar.

5.3.6.1 Ribosomal proteins

Group 1 identified all of the ribosomal proteins as a functionally similar group. It contained RPS2, RPL5, RPL13a, RPL21 and RPL27. It was the largest functionally similar groups of genes. RPS2 is a component of the 40S ribosomal subunit, and RPL5, RPL13a, RL21 and RPL27 are components of the 60S ribosomal subunit. It has been suggested however that several of the ribosomal proteins may have functions other than their role in protein synthesis. It has been shown for example that RPL13a can interact with the gamma activated inhibitor of translation (GAIT) element which is involved in the Cp translation silencing (Mazumder, Sampath et al. 2003; Chaudhuri, Vyas et al. 2007). RPS2 has been shown to interact with PRMT3, which has a role in inhibiting ubiquitination of RPS2. It has been suggested that this interaction may regulate the enzymatic activity of PRMT3 (Choi, Jung et al. 2008). RPL5, along with several other ribosomal proteins, have been shown to interact with Mdm2, inhibiting its E3 ligase activity and resulting in activation of p53 (Macias, Jin et al.).

5.3.6.2 Transcription factors

The second group identified by DAVID identified a group of transcription factors, and also placed the bait EGR1 in this functional group. This group contained ZMIZ2, TAF8 and NFAT5.

5.3.6.2.1 Nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5)

NFAT5 is a transcription factor of the nuclear factors of activated T cells family and is known to regulate gene expression after activation in response to osmotic stress. NFAT5 contains a conserved DNA-binding domain, the Rel-homology domain (RHD), which is also found in NF-κB (Macian, Lopez-Rodriguez et al. 2001; Muller and Rao 2010). NFAT5 is localised in the nucleus and requires dimerisation to bind to DNA. Upon activation, NFAT5 regulates the induction of inflammatory cytokines (Lopez-Rodriguez, Aramburu et al. 2001; Aramburu, Drews-Elger et al. 2006). NFAT5 is expressed in invasive breast cancer and it has been suggested that NFAT5 is involved in promoting tumour invasion in both human breast and colon cancer cells (Jauliac, Lopez-Rodriguez et al. 2002). S100A4, a member of the S100 calcium binding proteins, has been associated with the progression of several cancers including colon cancer and it has been demonstrated that hyperosmotic stress induces S100A4 expression, mediated via NFAT activation by Src, in colon cancer cells (Chen, Sastry et al. 2011).

5.3.6.2.2 TAF8 RNA polymerase II, TATA box binding protein (TBP)-associated factor (TAF8)

TAF8 is a TATA-binding protein (TBP)-associated factor (TAF). It is one of several TAFs that comprise a transcription factor complex TFIID which nucleates the transcription preinitiation complex (PIC) assembly with RNA polymerase II on protein-coding genes. TAF8 contains a histone-fold domain and has been shown to interact with several of the TFIID subunits (Guermah, Ge et al. 2003; Soutoglou, Demeny et al. 2005; Demeny, Soutoglou et al. 2007; Alvarez and Woolf 2011). TAF8 has been identified as having a role in the regulation of gene expression in prostate cancer (Huang, Beliakoff et al. 2005; Alvarez and Woolf 2011).

5.3.6.2.3 Zinc finger, MIZ-type containing 2 (ZMIZ2)

ZMIZ2 is a member of the PIAS (protein inhibitor of activated signal transducer and activator of transcription) family. ZMIZ2 is a transcriptional coregulator of the androgen receptor AR. It has been shown to interact with AR, which enhances AR-mediated transcription, which is further augmented by an interaction between ZMIZ2 and PIAS3 (Huang, Beliakoff et al. 2005; Beliakoff and Sun 2006; Peng, Lee et al. 2010). There is a well established link between AR and prostate cancer (Balk and Knudsen 2008). It has been shown that expression of AR is decrease in colorectal cancer tumours compared with normal mucosa in male patients only (Leon, Casado et al. 2011). This is a very interesting candidate as there is a higher incidence rate of CRC in men than women.

5.3.6.3 Translation factors

There are two proteins that were identified in the screen that are involved in the translation of proteins, eEF1A1 and eIF4A1

5.3.6.3.1 Eukaryotic translation initiation factor 4A1 (eIF4A1)

eIF4A1 is an ATP-dependent RNA helicase that is involved in the initiation phase of the protein translation process. Its function is to unwind RNA at the 5'-untranslated region (UTR) of mRNA (Kapp and Lorsch 2004). It forms part of a multi-protein complex that also contains eIF4G, eIF4E, eIF4B and eIF4H (Rozovsky, Butterworth et al. 2008; Marintchev, Edmonds et al. 2009).

5.3.6.3.2 Eukaryotic translation elongation factor 1 alpha1 (eEF1A1)

eEF1A1 is involved in the elongation phase of translation and is known to be involved in the transport of aminoacyl tRNAs to the ribosomal A-site. It forms a ternary complex with GTP and the aminoacyl tRNA (Kapp and Lorsch 2004; Lamberti, Caraglia et al. 2004). eEF1A1 is differentially expressed in CRC tumours (Bertucci, Salas et al. 2004).

5.3.6.4 Remaining proteins

5.3.6.4.1 Brain abundant, membrane attached signal protein 1 (BASP1)

BASP1 was identified as a membrane and cytoskeleton- associated protein. It has been shown to associate with WT1 and negatively regulate the transcriptional activity of WT1 (Carpenter, Hill et al. 2004; Green, Wagner et al. 2009) (Carpenter et al, 2004; Green et al 2008). WT1 is a zinc-finger protein that is related to EGR1 and is mutated in Wilms' tumour (Scharnhorst, Menke et al. 2000). It is also suggested to have a role in the regulation of actin dynamics (Zakharov and Mosevitsky 2010). BASP1 has been shown to be a target of copy number increases in cervical cancer through gain of chromosome 5q (Scotto, Narayan et al. 2008). Conversely BASP1 been shown to have decreased expression in hepatocellular carcinoma (HCC) and is aberrantly methylated in HCC tissue compared with non-tumour liver tissue (Moribe, Iizuka et al. 2008). A study by Hartl et al, 2009 identified BASP1 as a potential tumour suppressor, as expression of BASP1 blocks myc-induced cell transformation and prevents activation and repression of myc-regulated genes (Hartl, Nist et al. 2009).

5.3.6.4.2 Chromosome 11 open reading frame 17 (C11orf17/BCA3/AKIP1)

C11orf17 was characterised as a breast cancer associated gene (BCA3), with high mRNA expression in breast and prostate tumour cell lines (Kitching, Li et al. 2003). It is localised to the nucleus and has been shown to interact with protein kinase A (Sastri, Barraclough et al. 2005). It also interacts with p65 and is suggested to enhance NF-κB mediated transcription (Gao, Asamitsu et al. 2008).

5.3.6.4.3 C-type lectin domain family 2, member D (CLEC2D)

CLEC2D was identified to code for a lectin-like receptor (LLT1) expressed on natural killer (NK) cells (Boles, Barten et al. 1999). It has been suggested to play a role in immune regulation and was shown to induce IFN-γ in NK cells but does not result in cytotoxicity (Mathew, Chuang et al. 2004). CLEC2D has alternatively

spliced transcripts that code for several different protein isoforms (Germain, Bihl et al. 2010).

5.3.6.4.4 FK506 binding protein 9 (FKBP9)

FKBP9 was identified as a new member of a group of proteins (FKBP) which are known to bind to immunosuppressant drugs e.g. FK506, rapamycin and ciclosporin. It is highly expressed in the heart, muscle lung and kidney (Jo, Lyu et al. 2001). The FKBP proteins have several cellular functions including protein trafficking, transcription, protein folding and receptor signalling (Kang, Hong et al. 2008), but the exact function of FKBP9 is not yet known.

5.3.6.4.5 GNAS complex locus

GNAS is a complex imprinted locus and codes several gene products including the stimulatory G protein α-subunit ($G_s\alpha$) which mediates c-AMP production by activation of adenylate cyclase. It can encode $XL\alpha_s$ which is only expressed from the paternal allele (Xie, Plagge et al. 2006; Krechowec and Plagge 2008). GNAS has been shown to be differentially expressed in epithelial ovarian cancer patients (Tominaga, Tsuda et al. 2010).

5.3.6.4.6 Nascent polypeptide-associated complex alpha subunit (NACA)

NACA is a ribosome-associated protein that binds to nascent polypeptide chains as they leave the ribosome which prevents them from binding to cytosolic proteins (Wiedmann, Sakai et al. 1994). NACA has been proposed to have several other functions in the cell other than a role in protein translation. It has been suggested to have a role in the regulation of FADD function (Stilo, Liguoro et al. 2003), hypoxia-mediated apoptosis (Hotokezaka, van Leyen et al. 2009) and in the regulation of erythroid differentiation (Lopez, Stuhl et al. 2005).

5.3.6.4.7 Nucleophosmin (NPM1)

NPM1 is a phosphoprotein that is localised in the nucleolus, though it is known to rapidly translocate between the nucleus and the cytoplasm. It is thought to have many functions in the cell such as ribosome biogenesis, cell proliferation, stress-induced apoptosis and is a positive regulator of ARF (Gjerset 2006; Grisendi, Mecucci et al. 2006). NPM1 is differentially expressed in CRC tumours (Bertucci, Salas et al. 2004).

5.3.6.4.8 Polyhomeotic homolog 1 (PHC1)

PHC1 is a component of the Polycomb Repressive Complex 1, which is involved in gene silencing possibly through reorganisation of chromatin into a compact structure (Simon and Kingston 2009). PHC1 has been shown to be overexpressed in testicular germ-cell tumours as well as in pituitary adenomas. However expression of PHC1 was not detected in normal colon tissue or in colorectal adenocarcinomas (Rodriguez, Jafer et al. 2003; Sanchez-Beato, Sanchez et al. 2006).

5.3.6.4.9 Phosphoserine phosphatase (PSPH)

PSPH is a phosphotransferase that is involved in the hydrolysis of L-phosphoserine in a magnesium dependent manner that results in formation of L-serine (Collet, Stroobant et al. 1999; Wang, Cho et al. 2002). It has been shown that calcium inhibits the enzymatic activity of PSPH (Peeraer, Rabijns et al. 2002). It has also been suggested that PSPH may play a role in regulating the proliferation of neural stem cells (Nakano, Dougherty et al. 2007).

5.3.6.4.10 Parathymosin (PTMS)

PTMS is a small (11.5kDa) zinc binding protein. It is predominantly localised in the cytoplasm but does contain a functionally active nuclear localisation signal (Clinton, Graeve et al. 1991; Trompeter, Blankenburg et al. 1996) and has been shown to have a role in the inactivation of phosphofructokinase. PTMS has been shown to interact with histone H1 and is suggested to have a role in the modulation of H1 interaction with chromatin (Kondili, Tsolas et al. 1996; Martic, Karetsou et al. 2005).

5.3.6.4.11 Solute carrier family 7 (amino acid transporter, L-type), member 8 (SLC7A8)

SLC7A8 was identified as a heterometric amino acid transporter with SLC3A2, a glycoprotein heavy chain 4F2 and is highly expressed in the skeletal muscle, intestine, kidney and placenta, as well as the liver, heart and lung. It has broad substrate specificity for zwitterionic amino acids (Bassi, Sperandeo et al. 1999; Pineda, Fernandez et al. 1999; Rossier, Meier et al. 1999).

5.3.6.4.12 Yip1 interacting factor homolog A (YIF1A)

Not much is known about the function of this protein but it is thought to be localised in the Golgi apparatus (Jin, Zhang et al. 2005). It has been shown that YIF1A is significantly differentially expressed in normal and squamous cell carcinoma tumour samples by microarray analysis, with expression significantly higher in the tumour samples (Sugimoto, Seki et al. 2009).

5.3.7 Determination of false positives

Once the potential EGR1 protein-interactions had been identified, they were then screened for false positives. The 'prey' protein in the pGADT7 was isolated from the yeast colony and co-transformed with both an empty pGBKT7 vector and with a pGBKT7 vector containing EGR1. If the interaction is true it will grow only in the presence of the EGR1 protein and no colonies will be evident on the plate with transformed empty pGBKT7 and the prey. However if it is a false positive, the 'prey' protein will activate the transcription of the ADE2 and HIS3 reporter genes in the absence of the interacting protein indicating that the prey protein does not require an interaction to activate the reporters and colonies will be evident on the plate with transformed empty pGBKT7 and the prey plasmid.

Four proteins demonstrated slight auto-activation of the ADE2 and HIS3 reporters, TAF8, CLE2D, ZMIZ2 and RPL27. These results are not completely unexpected as ribosomal proteins are one of the main groups of proteins known to pulled out as false positives in yeast two hybrid screen, as are transcription factors, whose, as stated previously, activation domain can activate the gene transcription of the reporter genes in the absence of an interacting protein.

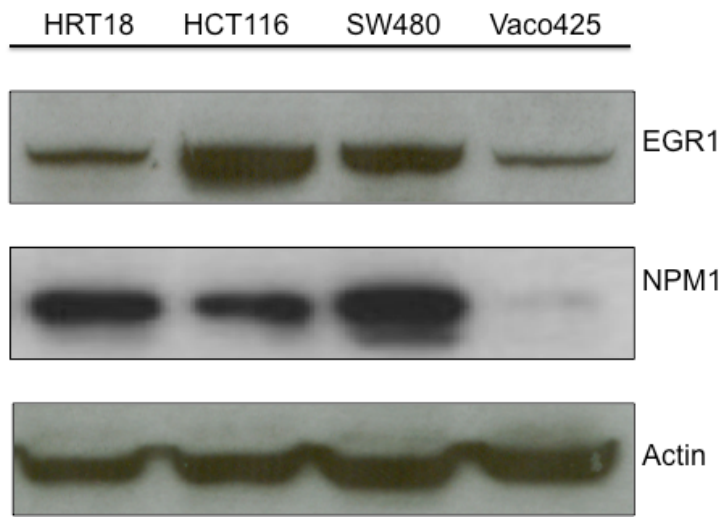
5.3.8 Investigation into the potential interaction of EGR1 and NPM1

The first potential interaction of EGR1 to be investigated was with the NPM1 protein, as it is has been implicated in tumorigenesis, and has been shown to be over-expressed CRC (Nozawa et al, 1996).

5.3.8.1 Expression of NPM1 in CRC cells

In order to confirm a potential interaction between EGR1 and NPM1, we first determined the expression and localisation of NPM1 in CRC cells. NPM1 was found to be expressed in all four of the CRC cell lines tested, with the lowest expression evident in Vaco425 (Figure 5.7 A) Although NPM1 is a nucleolar protein it is known to shuttle between the cytoplasm and the nucleus. Using a subcellular localisation kit, it is clear that in SW480 cells, NPM1 is localised in both the cytosol and the nucleus with faint localisation in the membrane/organelle fraction (Figure 5.7 B).

A.



B.

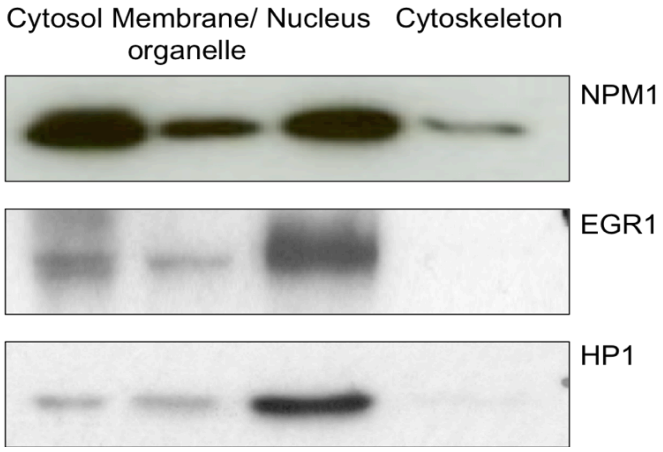


Figure 5.7: Expression and localisation of NPM1 in CRC cells

A. Protein extracts (30μg) were prepared from HRT18, HCT116, SW480 and VACO425 cells, resolved on a 10% acrylamide gel and probed with anti-NPM1 and anti-EGR1 antibodies. An anti-actin antibody was used as a loading control. **B.** SW480 cells were extracted into four cellular compartments, 30μg of protein was resolved on a 10% acrylamide gel and probed with anti-NPM1, anti-EGR1 and anti-HP1 antibodies. Both membranes have been stripped and re-probed with the antibodies shown here.

5.3.8.2 Immunoprecipitation with EGR1 and NPM1

The interaction between NPM1 and EGR1 proteins was investigated by immunoprecipitation. The experiments were conducted using the experimental conditions for immunoprecipitation as determined in the previous chapter (Chapter 4.3.3.3) The immunopreciptation was conducted using agarose G beads, and lysis buffer 4 that contains 300mM NaCL and 1% Triton x100. The interaction between EGR1 and NPM1 was investigated in both SW480 and HRT18 cells. Figure 5.8 demonstrates that there is a weak interaction between EGR1 and NPM1 in both HRT18 and SW480 cells, with and without LPS treatment. As in previous experiments there is an increase in non-specific background in the IgG mouse control after LPS treatment.

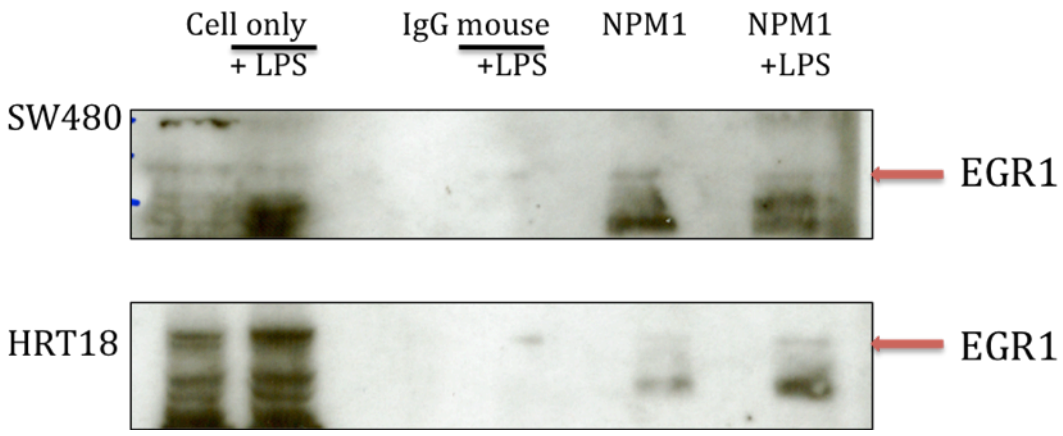


Figure 5.8: Immunoprecipitation with NPM1 and EGR1

Cells were treated with 1μg LPS for 24 hours, extracted using lysis buffer 4 and immunoprecipitated using agarose G beads and anti-NPM1. IgG goat was used as a negative control. Immunoprecipitated complexes were run on a 10% resolving gel and probed with anti-EGR1.

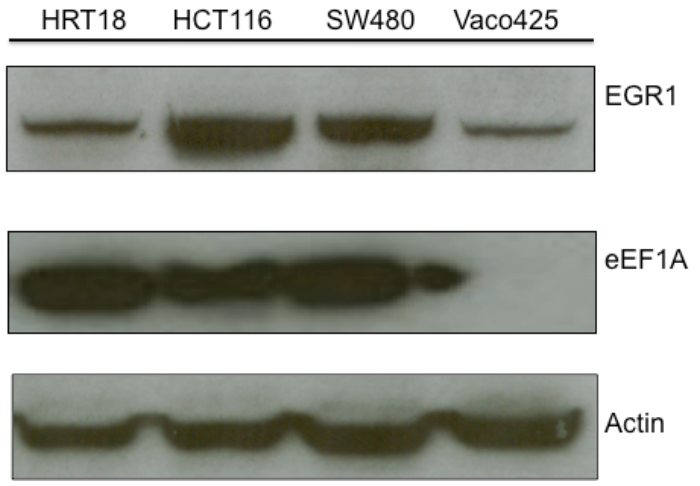
5.3.9 Investigation into the potential interaction of EGR1 and eEIF1A1

The second protein interaction to be investigated from the Y2H screen is eEIF1A1, which is, as previously stated, involved in the elongation stage of protein translation. It has been implicated in tumorigenesis and is over-expressed in colon cancer (Zhang, Zhou et al. 1997; Bertucci, Salas et al. 2004; Lamberti, Caraglia et al. 2004).

5.3.9.1 Expression of eEIF1A1 in CRC cells

As before, we first confirmed that eEIF1A1 is expressed in CRC cells. eEIF1A1 is strongly expressed in HRT18, HCT116 and SW480 cell lines (Figure 5.9.A). Expression of eEF1A is evident in both the cytosol and nuclear fraction of SW480 cells (Figure 5.9.B).

A.



B.

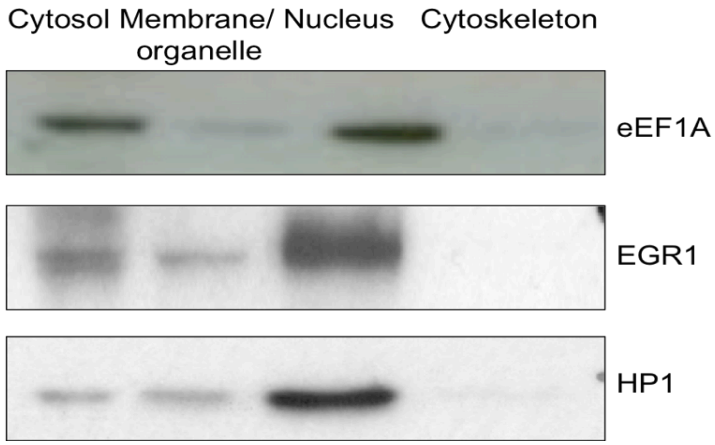


Figure 5.9: Expression and localisation of eEF1A1 in CRC cells

A. Protein extracts (30μg) were prepared from HRT18, HCT116, SW480 and Vaco425 cells, resolved on a 10% acrylamide gel and probed with anti-eEF1A1 and anti-EGR1 antibodies. An anti-actin antibody was used as a loading control**. B.** SW480 cells were extracted into four cellular compartments, 30μg of protein was resolved on a 10% acrylamide gel and probed with anti-eEF1A1, anti-EGR1 and anti-HP1 antibodies. Both membranes have been stripped and re-probed with antibodies shown here.

5.3.9.2 Immunoprecipitation with EGR1 and eEF1A1

The interaction between eEF1A1 and EGR1 was investigated by immunoprecipitation. The experiments were conducted using the previously developed experimental conditions for immunoprecipitation as before. The interaction between EGR1 and eEF1A1 was investigated in both SW480 and HRT18 cells. However there is only a faint interaction detected between EGR1 and eEFIA1 in HRT18 cells (Figure 5.10). Although EGR1 has lower level of expression in HRT18 and that the original yeast two hybrid was conducted using an SW480 cell line library, an interaction was not detected in SW480 cells using this method. Interestingly the interaction between EGR1 and eEF1A1 appears to be decreased in the cells after treatment with LPS.

There is > 92% homology between eEF1A1 and eEF1A2 (Soares, Barlow et al. 2009) and as most antibodies do not differentiate between the two isoforms, it cannot be said that the interaction detected is with the eEF1A1 isoform. Antibodies specific to the two isoforms became available to use (from Cathy Abbott, MMC), however no interaction was detected with either the A1 or A2 isoform in either HRT18 or SW480 cells. As these antibodies consist of a small specific peptide to each isoform, the experiment was conducted by pulling down with an EGR1 anti-mouse antibody (not previously used in IP experiments) and probing with the A1 and A2 anti-sheep antibodies. It therefore may be necessary to further alter the experimental conditions in order to detect a stronger interaction or to confirm the interaction with the A1 isoform and in SW480 cells also.

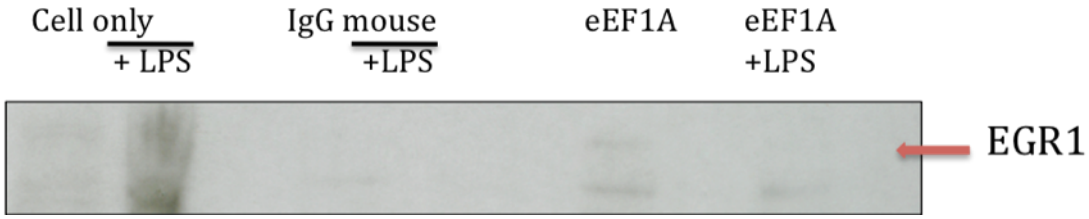


Figure 5.10: Immunoprecipitation with EGR1 and eEF1A1

Cells were treated with 1μg LPS for 24 hours, extracted using lysis buffer 4 and immunoprecipitated using agarose G beads and anti-NPM1. IgG goat was used as a negative control. Immunoprecipitated complexes were run on a 10% resolving gel and probed with anti-EGR1.

5.4 Discussion

A yeast-two hybrid screen conducted using EGR1 as bait identified 22 novel interacting proteins of EGR1 in a SW480 epithelial cell library. As transcription factors can activate transcription of the reporters in the yeast without an interacting protein, an EGR1 construct without its transactivation domain as well as a full length EGR1 construct was cloned into the pGBKT7 vector. However there was no evidence of auto-activation by the EGR1 construct with either the full length or the truncated EGR1. The yeast-two hybrid screen was conducted using an SW480 CRC cell line library, to further increase our knowledge of EGR1 interacting proteins in epithelial CRC cells. Interestingly all of the proteins indentified were detected using full length EGR1, with only two proteins being identified as also interacting with EGR1ΔAct but gives valuable information with regards to the possible EGR1 interacting domain for these proteins.

In total ~ 800 colonies were selected and tested for confirmation of interaction using the colon lift filter assay. Of these, approximately half of all colonies selected were false positives and disgarded. Of the remaining, these were sequenced and analysed to see if the colony contained a protein sequence of the human gene. Only sequences that coded for a protein were selected. In total 22 protein sequences were isolated, however 4 of these were found to be false positives, that is proteins whose structure allows them to activated gene transcription in the Y2H system in the absence of an interacting protein. It should be noted however that due to time constraints only a small scale Y2H screen was conducted and there was not the opportunity to screen many more interacting colonies. It is therefore possible that many more proteins could be determined as interacting with both EGR1Fl and EGR1ΔAct if a larger scale screen was conducted. However this small screen provided numerous possible targets for further investigation

There are several limitations of the yeast-two method that should also be taken into account. As the yeast-two hybrid screen requires translocation of the proteins into the nucleus to activate the reporter gene expression this method can therefore not detect

interactions with membrane proteins or membrane-bound proteins. The proteins of interest must be able to fold correctly and be stably expressed in yeast, and it is possible that fusion of the DNA-binding domain or activation domain can change the confirmation of either the prey or the bait, which may have an impact on binding sites and interactions. It is also necessary to confirm any interaction by another method in your cell type of choice, such as by immunoprecipitation, therefore it should be kept in mind that of the 18 positive interactions determined by the screen, not all of these proteins may interact with EGR1 outside of the yeast system.

Bearing in mind these limitations, the yeast two-hybrid screen is still a powerful tool to detect novel protein interaction and in this screen many potential interactions and proteins of interest were detected. It was decided to first investigate the potential interaction between EGR1 and NPM1. NPM1 is currently of interest in the CCGG group in relation to aspirin mediated apoptosis of CRC cells. As stated, NPM1 is a phosphoprotein that is mainly localised in the nucleolus but can rapidly shuttle between the nucleus and the cytoplasm. A mutant form of NPM1 commonly found in acute myeloid leukemia (AML) is aberrantly localised in the cytoplasm (Falini, Mecucci et al. 2005). NPM1 is also known to translocate from the nucleolus in response to genotoxic stress and UV irradiation (Rubbi and Milner 2003). As such it is evident that NPM1 can localise in the nucleus or the cytoplasm and therefore it is unknown where EGR1 and NPM1 are likely to interact in the cell, as EGR1 can also localise in the cytoplasm or the nucleus. Indeed previous work in the CCGG has indicated that EGR1 may localise to the nucleolus and as such cannot be ruled out as a cellular compartment for the interaction.

NPM1 has been shown to have a variety of cellular functions. It was initially identified as having a role in ribosome biogenesis and contributing to cell growth. It is known to bind to p53, ARF and p21 and, like EGR1, appears to act as both a tumour suppressor and tumour promoter, depending on the circumstances (Gjerset 2006; Grisendi, Mecucci et al. 2006; Xiao, Zhang et al. 2009). NPM1 has been shown to activate p53 transcriptional activity (Gjerset 2006). NPM1 is over-expressed in various tumours, including CRC (Nozawa, Van Belzen et al. 1996). Over-expression of NPM1 has been shown to lead to increased cell growth, as well

as inhibiting pro-apoptotic pathways (Grisendi, Mecucci et al 2006). Interestingly NPM1 has recently been shown to interact with the kinesin-like protein Eg5 and has a role in modulating its ability to destabilise microtubules (Wang, Gao et al. 2010). A microarray study demonstrated that there is an increase in NPM1 expression in CRC tumours (Bertucci, Salas et al. 2004).

It is unclear what effect an interaction between NPM1 and EGR1 may have on the function of either protein. However the cellular functions of both EGR1 and NPM1 are varied and depend on cellular conditions such as stress and therefore it may be difficult to fully elucidate the functional effects of an interaction. Initially it would be interesting to determine if an interaction between EGR1 and NPM1 has any effect on the transcriptional activity of EGR1, using the luciferase gene reporter assays that are available for both EGR1 and NPM1 as a resource in the lab, followed by experiments to determine if loss of NPM1 in CRC cells had an effect on the downstream targets of EGR1. It would be necessary however to conduct these experiments using NPM1 siRNA as NPM1 knockout mice are embryonic lethal (Grisendi, Bernardi et al. 2005). A further area of research would be to determine if loss of NPM1 has any effect on the localisation of EGR1 in the cell, as NPM1 has been reported to act as a molecular chaperone (Kerr, Birse-Archbold et al. 2007), and as it also has links with cytoskeleton organisation, it would be interesting to determine if its loss has any effect on EGR1s interaction with the cytoskeleton. Also as NPM1 responds to stress by translocation to the cytoplasm investigating the interaction between EGR1 and NPM1 under cellular stress conditions such as UV or LPS would aid in understanding their possible role in carcinogenesis.

The second protein investigated was the translation elongation protein eEF1A1. Although the function of eEF1A1 in the initation of the elongation stage of protein translation is well characterised, it is also thought to have a role in several other cellular processes. eEFIA1 can form a complex with both tubulin and actin and it has been shown to have an effect on cytoskeletal organisation (Condeelis 1995), which is intriguing given that we have shown an interaction with EGR1 and tubulin in CRC cells. Interestingly an oncogene, PT-1 (tumour inducing gene) was identified in prostate carcinoma which appears to consist of a mutated eEF1A1 protein fused with

a region of 5' UTR that has a high homology (~85%) to a 23S ribosomal gene from *Mycoplasma hyponeumoniae* (Gopalkrishnan, Su et al. 1999). PT-I is also expressed in breast, colon, and lung cancer cells, which have differential expression of EGR1. eEF1A has been reported to be involved in stress-induced apoptosis (Lamberti, Caraglia et al. 2004) and been shown to be differentially expressed in CRC by microarray, with higher levels of eEF1A1 evident in CRC tumours (Bertucci, Salas et al. 2004).

There are two isoforms of eEF1A, eEF1A1 and eEF1A2, which are encoded by different genes but are nearly identical at the protein level (Soares, Barlow et al. 2009). The A1 isoform is ubiquitously expressed whereas the A2 protein expression is limited to the brain, heart and skeletal muscle (Lee, Francoeur et al. 1992). We have shown that eEF1A is expressed in CRC cells, and there is a faint interaction evident in HRT18 cells. Further optimisation may be required to detect a stronger interaction in HRT18 and in SW480 cells. Interestingly the interaction that was detected in HRT18 cells was reduced when using LPS treated cells. Further experiments could involve determining if this reduction is only evident after treatment with LPS, or if other inflammatory or stress stimuli cause a similar reduction or alteration.

As previously stated the antibody used for the immunoprecipitation is known to be non-specific in that it detects both isoforms eEF1A1 and eEF1A2. Therefore further work is necessary to confirm which isoform it is that is interacting with EGR1, indeed it may be that both isoforms do. As with NPM1, it is unclear as to the effects of an interaction between EGR1 and eEF1A. eEF1A1 has been shown to interact with other proteins outside of its role as a translation factor, some of which are regulated by phosphorylation (Ejiri 2002) and as phosphorylation has also been shown to play a role in some EGR1 protein interaction (Adamson, Yu et al. 2005), it is possible that the interaction between EGR1 and eEF1A1 is similarly regulated or plays a role in the effect of the interaction.

6 Conclusion

EGR1 has long been of interest due to its role as a transcription factor involved in the regulation of several different pathways in both cancer and inflammation. When it was first identified in 1988, it was shown to have a role in the regulation of cell growth and differentiation. Since then EGR1 has been shown to be induced by many different growth and stress stimuli, resulting in downstream activation of targets involved in cell growth and differentiation and the inflammatory response pathways. EGR1 became an interesting candidate for a role in carcinogenesis, not only because of its ability to effect cell growth, but due to the fact that it is differentially expressed in many different cancers, including breast and prostate. The role of EGR1 has been investigated predominantly in prostate cancer, but it has also been shown to be differentially expressed in colorectal cancer. When studies conducted by the CCGG and GI unit found that *EGR1* is downregulated in the tumour mucosa of patients with CRC compared with normal colonic mucosa (in silico work), as well as demonstrating an association between common variants of *EGR1* with disease phenotype in both CRC and IBD it became an interesting candidate gene for both CRC and IBD.

Whilst previous studies have looked at the expression of *EGR1* in CRC patients using microarray and qRT-PCR, this is the first study to look at the expression of *EGR1* in matched normal and tumour samples. This study used 30 matched normal mucosa and tumour samples to determine if there is any differential expression of *EGR1* in CRC patients. This study is also the first to determine if there is any correlation with genotype of three *EGR1* SNPs with its expression in CRC as well as determining if the differential expression of *EGR1* observed is associated with aberrant methylation of the *EGR1* promoter in epithelial cells.

We have shown that *EGR1* is differentially expressed in CRC patients. There appears to be three groups of patients as regards *EGR1* expression. One group of patients have significantly higher levels of *EGR1* expression in the tumour compared with the normal mucosa, another group of patients has significantly lower *EGR1* expression in the tumour compared with the normal mucosa whilst several patients show no

difference in expression between the normal mucosa and tumour. A large region of the *EGR1* promoter was investigated for methylation of the CpG islands, however no methylation was detected in the large region examined. No association with *EGR1* expression and the genotype of the *EGR1* SNPs was proven, however a much greater sample size would be necessary to thoroughly examine this and any possible haplotype effect.

A study by Habermann et al. has demonstrated that *EGR1* is differentially expressed at different stages in CRC development, with differential expression seen between normal mucosa and adenoma, and between adenoma and carcinoma (Habermann, Paulsen et al. 2007). They observed a significant decrease in expression between the normal mucosa and adenoma, with a significant increase in EGR1 between adenoma and carcinoma tissue. This data agrees with an analysis we conducted using Oncomine which showed a study by Skrzypczak et al. demonstrated a decrease in *EGR1* in colorectal adenocarcinoma with a study conducted by Hong et al. observing that EGR1 was differentially expressed in colon and colorectal carcinoma, with higher *EGR1* expression seen in the colorectal carcinoma (Hong, Downey et al. 2010; Skrzypczak, Goryca et al. 2010). Therefore it is necessary to determine if the differential expression between matched normal and tumour seen in this study can is associated with tumour stage of the patients, as this would allow a greater understanding to the pattern of differential expression of *EGR1* in CRC and its role in carcinogenesis.

It is possible that the differential levels of EGR1 seen in this study in the normal and tumour may have very different effect on activation of its downstream targets and further analysis would allow us to determine if EGR1 has a greater impact on tumour suppressor genes or oncogenes, ie whether is EGR1 is growth-promoting or inhibiting in this cancer. CRC is a heterogeneous disease so it is possible that EGR1 may have a different role in different tumours, or at the different stages of the tumour progression. The studies mentioned above have indicated that EGR1 expression is lower in the adenoma stage of CRC with an increase in the carcinoma stage. If EGR1 expression is higher at this later stage of the cancer it may have a role in activation of growth-promoting genes and metastasis genes. This is the stage of tumour

progression in which p53 mutation usually occurs, and it has been shown that EGR1 can regulate the expression of p53 (Yu, Baron et al. 2007). It has also been observed that the mutant status of p53 has an effect on the function of EGR1 (Weisz, Zalcenstein et al. 2004).

To investigate mechanisms by which *EGR1* was differently expressed in CRC patients the *KRAS* and *BRAF* mutation status were examined in these patients, an important pathway to investigate given its role in CRC and that *EGR1* expression is induced via the MAPK signalling pathway. This pathway also has an important role to play as a biomarker for the response to EGFR-targeted drugs for CRC, as mutations in *KRAS* are a good indicator for the effectiveness of the response of cetuximab and panitumumab (Soulieres, Greer et al. 2010). At the adenoma stage of tumour progression EGR1 has been shown to be decreased, with correlates with occurrence of mutation of KRAS and BRAF, which are both involved in the MAPK pathway which regulates the expression of EGR1, with an increase of EGR1 in the carcinoma stage. In this study we observe that patients with KRAS or BRAF mutations had significantly higher levels of EGR1 in the tumour compared with the normal. It may be that mutations in KRAS/BRAF are associated with higher EGR1 expression. Given the role of APC in CRC and its proximity to *EGR1* on chromosome 5, it is necessary to determine if there is any correlation with APC mutation status and *EGR1* as well as looking to see if the differential expression is correlated with any other phenotypes, or disease characteristics such as stage of disease or site of tumour.

Given that mutations in CRC occur at different stages of the cancer progression, further analysis the mutation status of some of these genes would give an indication of what stage of cancer these patients were at, in the absence of patient data. We know the KRAS/BRAF mutation status of the patients, but analysing the samples for other mutations, p53 status for example, may indicate if any of the patients were in the final stages of tumour development, as presence of p53 mutation would indicate a late stage. It would be preferable to look at these data, along with APC, KRAS/BRAF and PTEN mutation status, with respect to clinical and pathological

data, which was not available in this study and to show that the increase in EGR1 in the tumour is around the carcinoma stage, with decrease in EGR1 in the tumour correlating with adenocarcinoma as the Oncomine analysis predicts. Given that EGR1 is known to regulate the expression of the tumour suppressor genes such as p53 and PTEN, an expression analysis of these downstream targets of EGR1, along with their mutation status, could provide an indication into the effect that the differential expression of EGR1 is having on key tumour suppressor genes. Again, despite the lack of patient data, a study to test for MSI instability would show if the patient samples are MSI or MSS, and given than MSI and CIN are usually mutually exclusive this would allow us to determine the type of genetic instability present in the patients which may provide further information about the type of CRC and *EGR1* expression.

This study also investigates the differential expression of *EGR1* mRNA in non-inflamed tissue of both UC and CD patients, comparing the expression with a group of healthy controls. We have also looked at the expression of *EGR1* in these three patients groups after treatment with inflammatory mediators, LPS, MDP, PGN and TNF. Again we examined a large region of the *EGR1* promoter to determine if aberrant methylation is associated with the differential expression, as well as looking at the genotype of the three *EGR1* SNPs. The expression of the *NAB2* inhibitor of *EGR1* was also examined to determine if there is any differential expression of it in the healthy controls, UC or CD patients.

This study shows that expression of EGR1 is decreased in the uninflammed mucosa of ulcerative patients and is significantly lower in Crohn's disease patients when compared with healthy controls. We determined that this decrease in expression is not associated with aberrant methylation of a large region of the EGR1 promoter, with the genotype of three *EGR1* SNPs or with differential expression of the *EGR1* inhibitor *NAB2*. Upon investigation of *EGR1* expression in the three patient groups after treatment with the inflammatory mediators we showed that as expected *EGR1* expression was significantly induced in the healthy controls after treatment with LPS and TNF. However we have demonstrated that the induction of *EGR1* in both UC and CD samples does not appear to act in the same manner as in the healthy controls.

In UC patients, LPS fails to significantly induce *EGR1* expression and treatment with PGN appears to reduce expression. In CD patients there is a reduction of *EGR1* after TNF treatment, and a significant reduction after LPS, demonstrating that in general IBD patients appear to have an aberrant induction of *EGR1* in response to inflammatory mediators.

One important area of further study would be to determine if there is differential expression of the EGR1 protein in both the CRC and the IBD patients, and indeed if there is any correlation at the protein level with the three EGR1 SNPs in these patients. There does appear to be differential expression of the EGR1 mRNA and protein in some of the CRC cells lines. In HRT18 and SW480 the mRNA levels (low and high respectively) do correlate with the protein levels but this is not the case in HCT116 (low mRNA levels, higher protein) and Vaco425 (high mRNA and low protein). It appears in the cell lines that presence of the EGR1 variants does result in higher EGR1 mRNA expression, with SW480 and Vaco425, both homozygous variant, having the highest levels of *EGR1* expression in the CRC cell lines but this does not correlate entirely with the protein expression but other factors might be involved in protein turnover. It would be necessary to undertake a much larger study to determine if there is any association with EGR1 expression and the EGR1 variants, but equally it appears that there may be differential expression or localisation the EGR1 protein also which should be investigated, as well as the response of the EGR1 protein to inflammatory mediators in IBD.

We have shown that the expression of EGR1 is decrease in patients with IBD in un-inflamed tissue. It appears that IBD patients with decreased EGR1 expression in the mucosa do not respond as expected to inflammatory mediators. This result was unexpected, as it was presumed that EGR1 would have high expression in these patients, thereby activating gene expression of the inflammatory genes that is regulates such as TF and PGE_2. However it may be that the lower levels of EGR1 in these patients has an effect on the gene expression in the cells. As EGR1 can activate the expression of both pro-and anti-inflammatory genes perhaps the loss of EGR1 has shifted the balance in favour of the pro-inflammatory genes. We have shown that EGR1 in these patients does not respond to inflammatory mediators as expected. In

normal cells there is a high expression of EGR1 in the mucosa. After exposure to LPS and TNF, the expression of EGR1 is induced. We do not know what genes EGR1 goes on to activate in the healthy patients in response to this activation by LPS and TNF, which would be an important next step in this analysis. EGR1 expression is lower IBD patient mucosa and it does not respond to LPS and TNF in the same manner as the healthy controls. If we could determine if the downstream activation of EGR1 genes in the healthy controls is different in the healthy controls, before and after LPS/TNF treatment, it may be that EGR1 does not activate the same gene expression in the IBD patients in response to LPS/TNF treatment. We do not yet know if this loss of EGR1 and its lack or response in IBD is a consequence of the disruption of the inflammation pathway in these patients or EGR1 has a role to play in the disruption of the pathway. Determination of the downstream gene activation of EGR1 in these patients may help to answer that question. Given that EGR1 has been shown to interact with NOD2 in CD patients, it may well have a role to play in CD. Elucidating the function of this interaction may help to determine if the interaction is itself playing a role in the development or progression of the disease.

One important area of further study would be to determine if there is differential expression of the EGR1 protein in both the CRC and the IBD patients, and indeed if there is any correlation at the protein level with the three EGR1 SNPs in these patients. There does appear to be differential expression of the EGR1 mRNA and protein in some of the CRC cells lines. In HRT18 and SW480 the mRNA levels (low and high respectively) do correlate with the protein levels but this is not the case in HCT116 (low mRNA levels, higher protein) and Vaco425 (high mRNA and low protein). It appears in the cell lines that presence of the EGR1 variants does result in higher EGR1 mRNA expression, with SW480 and Vaco425, both homozygous variant, having the highest levels of *EGR1* expression in the CRC cell lines but this does not correlate entirely with the protein expression but other factors might be involved in protein turnover. It would be necessary to undertake a much larger study to determine if there is any association with EGR1 expression and the EGR1 variants, but equally it appears that there may be differential expression or localisation the EGR1 protein also which should be investigated, as well as the response of the EGR1 protein to inflammatory mediators in IBD.

The expression level of EGR1 in a cell may be important in deciding its role in the cell. We still do not know if the differential expression of EGR1 mRNA in both CRC and IBD has an impact on its protein expression. This would be an important step in determining the role, if any, of the EGR1/NOD2 interaction in IBD. A greater knowledge of the other EGR1-protein interactions may also be useful in understanding its role in both CRC and IBD. The aberrant response of *EGR1* in IBD shown in this study and a future investigation into the EGR1 protein expression and its response to inflammatory mediators would be especially informative given that we have demonstrated an interaction between EGR1 and the CD susceptibility gene product NOD2. This study has investigated the expression and localisation of EGR1 and NOD2 in CRC cells and shown that NOD2 is predominantly localised in the cytoplasm and EGR1 is localised in both the nucleus and the cytoplasm, with the two proteins co-localising in the cytoplasm. We have demonstrated a positive interaction between EGR1 and NOD2 in SW480 cells, in untreated and LPS-treated cells. The confirmation that EGR1 and NOD2 do interact in CRC cells offers the potential to further examine this interaction in order to determine if it mediates any gene activation of target genes of both EGR1 or NOD2, or if the interaction facilitates a role outside that of a transcription factor. This finding further opens up the possibility that EGR1 has a significant role to play in IBD.

A limitation of this study has been the difficulty in acquiring reliable antibodies for EGR1 and NOD2. There are several commercially available antibodies for EGR1 which were tested and optimised in this thesis, however the availability of NOD2 antibodies is poor, and therefore it was not possible to do any IP experiments using EGR1 as a pull down and probing with NOD2. Further experiments characterising the interaction between NOD2 and EGR1 may be dependent on acquiring a suitable NOD2 antibody. However there are several key experiments that need to be conducted to further characterise the EGR1/NOD2 interaction. The first would be to determine the functional consequences, if any, of this interaction in the cell. Luciferase reporter assays using gene constructs with EGR1 promoter regions such as TF in conjunction with siRNA could be utilised to determine if NOD2 is required

for EGR1-mediated transcription, especially of genes involved in inflammation. Similarly reporter constructs with gene promoters of genes activated by NOD2 such as NF-κB could be utilised to determine if EGR1 has any role to play in NOD2-mediated gene transcription, with and without stimulation by MDP. Given that NOD2 localisation to the membrane has been determined to play a role in its MDP-induced activation of NF-κB, with the cytoskeleton thought to play a role in NOD2-mediated transcription, a second area to research would be to investigate the role of the cytoskeleton in the EGR1/NOD2 interaction, especially given that we have demonstrated an interaction with EGR1 and the tubulin cytoskeleton. In order to confirm that the cytoskeleton has a role to play in the interaction it would be necessary to determine what effect disruption or stabilising the cytoskeleton would have on the interaction between NOD2 and EGR1.

A third area of study would be to investigate whether EGR1 interacts only with wild-type NOD2, or whether it also interacts with any of the three NOD2 mutants. There are three common NOD2 mutations that have been found in CD which do not appear to respond to stimulation by MDP with the Leu1107fsinsC mutant unable to localise to the cell membrane. The three NOD2 mutants are available as a resource in the lab, gifted by Colin Stevens, as HA-tagged constructs. Experiments to determine the localisation of the NOD2 mutants in CRC cells, and whether they co-localised with EGR1 would lend to a further understanding of a role of the EGR1/NOD2 interaction, as would confirming if EGR1 physically interacts with any of the mutants in CRC cells.

By disrupting regions of EGR1 or NODD2 by site-directed mutatgenesis we could determine the regions of the protein that are involved in the EGR1/NOD2 interaction. There are three domains in EGR1, the zinc-finger DNA binding domain, a transactivation domain and a regulatory domain which is known to bind the two transcriptional co-factors NAB1/2, which have a predominantly inhibitory effect on the transcription of EGR1. Knowledge of the region of EGR1 that binds to the NOD2 may lead to a better understanding of role of the interaction, whether NOD2 may act as a co-activator of EGR1 transcription or have an inhibitory effect of EGR1. Similarly it would be useful to determine what region of NOD2 is involved in the

interaction with EGR1. The NOD2 mutations found in CD are localised in the LRR domain of NOD2 and might prove useful in determining the interacting domain on NOD2.

EGR1 is known to localise with components of the cytoskeleton in prostate cancer and it has been suggested that this localisation is necessary for the translocation of EGR1 from the cytoplasm to the nucleus in these cells. We have shown that EGR1 also co-localises and interacts with components of the cytoskeleton in CRC cells, and it may be that this interaction is required for its translocation in this type of cancer cell also. We have demonstrated that EGR1 is localised in both the nucleus and the cytoplasm in CRC cells, and indeed there appears to be EGR1 present in the membrane bound and organelle fraction which may have other implications for its interaction with NOD2. Experiments investigating if the interaction between EGR1 and tubulin is involved in the translocation of EGR1 between the cytoplasm and nucleus would determine if the role of this interaction is similar in CRC as in prostate cells, and involved in the regulation of EGR1 translocation and localisation within the cell especially given that SW480 and HRT18 cells appear to have differences in the cellular localisation of EGR1, with more cytoplasmic EGR1 evident in HRT18 cells and greater nuclear localisation of EGR1 in SW480. Investigating the EGR1 and cytoskeletal interactions during cell cycle arrest, or at different stages of the cell cycle could determine if the localisation of EGR1 in the nucleus or the cytoplasm oscilliates during the cell cycle and whether it correlates with its interaction with tubulin.

Although we have demonstrated an interaction with EGR1 and NOD2, and EGR1 and tubulin in SW480 cells in the presence and absence of LPS treatment, further experiments into the interaction in the presence of stress stimuli and inflammatory mediators needs to be conducted, especially given the abnormal response of *EGR1* mRNA after treatment with the inflammatory mediators in CD patients as demonstrated by the qRT-PCR experiments. Also given how stimulation of EGR1 by different stress stimuli not only effects the post-translational modifications of EGR1 but also its interactions with other transcription factors such as p300 and CBP with different effects on the targets of EGR1, further investigation into the interaction

between EGR1 and NOD2 is warranted under different stress conditions to see if there is any difference in the interaction, post-translational modifications of the proteins and their functions in the cell. A key experimemt would be to determine if EGR1 and NOD 2 interaction occurs using a phospho-EGR1 antibody, and whether under different stress conditions if this interaction is gained or lost.

Little is known about proteins that may interact with EGR1. One of the first proteins found to interact with EGR1 acts as an inhibitor of EGR1. Other proteins that interact with EGR1 are also thought to be involved in either its regulation or its role as a transcription factor. Although there are several reports of yeast-two hybrid screen conducted using EGR1 as a bait protein, this is the first time it has been reported in epithelial cell lines. The yeast-two hybrid screen was conducted using two EGR1 constructs, a full length EGR1 and an EGR1 construct minus its transactivation domain. The screen was conducted against an SW480 cell line library.

This study identified 22 novel potential interacting proteins of EGR1, which are involved in several different cellular functions. It is of great interest to further investigate these protein interactions, where they occur, and if they have an important cellular role. Two interesting candidate proteins were selected for further investigation, NPM1 and eEF1A1. NPM1 is a good candidate gene for CRC, as it has been shown to be highly expressed in CRC tumour compared with normal tissue, as has our other tested candidate, eEF1A1.

These novel interactions were confirmed in CRC cells, as well as preliminary studies to determine their localisation within the cells to determine if they would be in proximity to EGR1 to allow an interaction to occur. We have confirmed that both NPM1 and eEF1A1 are expressed in CRC cells, and have a positive interaction with EGR1. In the case of NPM1 this interaction has been demonstrated in SW480 and HRT18 cells, with and without treatment of LPS, however an interaction with eEF1A1 has so far only been detected in HRT18 in untreated cells. Like the interaction with NOD2, it does appear that the interaction with NPM1 is likely to occur in the cytoplasm of the CRC cell lines, but it may turn out that the interaction

is localised in the nucleus and be involved in the transcritptional activity of either NPM1 or EGR1. Previous work in the CCGG has demonstrated that EGR1 can also localise to the nucleolus in CRC cells by immunocytochemistry and this is a potential site of interaction with EGR1 for both NPM1 and eEF1A1.

The confirmation that EGR1 does interact with both NPM1 and eEF1A1 opens up many different areas of research to be considered. As with the interaction with NOD2 it is necessary to elucidate the function of an interaction between EGR1 and NPM1 or eEF1A1. As regards the interaction with NPM1 there are several resources available in the CCGG that would allow for further investigation into the interaction with EGR1 including several different commercial antibodies, luciferase reporter assays for genes that are regulated by NPM1 and NPM1 siRNA primers. However given the variety of cellular roles that have been attributed to NPM1, fully determining its role with EGR1 may prove difficult.

As discussed there are two isoforms of eEF1A found, eEF1A1 and A2, which are very closely related. It would be necessary to confirm that it is the A1 isoform that interacts with EGR1 and to either determine or rule out an interaction with the A2 isoform. The commercially available antibodies are not specific enough to determine which isoform is involved in the interaction we have detected with EGR1 but specific antibodies have been made available (gift from Cathy Abbott). Due to the nature of these antibodies it would first be necessary to optimise the experiments to allow pull down with EGR1 and use the eEF1A antibodies to probe to determine an interaction before investigating the role that this interaction may have in relation to the function of EGR1 or eEF1A in the cell. However this interaction and the finding that EGR1 potentially interacts with many of the candidate proteins determined in the yeast-two hybrid screen opens that possibility of novel roles/pathways for EGR1 that may be important in furthering our understanding of the role of EGR1 in cancer and inflammation.

One final area of future research in relation to the various novel interacting proteins of EGR1 found in thesis would be to determine if EGR1 interacts with these proteins

in colon tissue of patients which could be achieved by conducting the interaction experiments using protein extracts from the normal mucosa and tumour of CRC patients, and particularly in the case of the EGR1/NOD2, in IBD patients to determine if the EGR1 interactions occur in cancer and/or in IBD. This would help us to begin to determine if the EGR1/NOD2 interaction has a functional role in IBD, and indeed if any of the novel proteins found in the yeast-two hybrid screen have a role in CRC/IBD with EGR1. Experiments are also being conducted in the CCGG to establish an ex-vivo model of both normal mucosa and tumour, which would present a further resource in looking at how EGR1 is expressed and localised and responds to stress stimuli in the colon over a time course as well as investigating the various EGR1-protein interactions.

To conclude, this study has shown that *EGR1* is differentially expressed in matched normal mucosa and tumours of CRC patients, as well as in patients with IBD. We have found no correlation between this differential expression and the genotype of three *EGR1* SNPs or the methylation of a large region of the *EGR1* promoter. In the case of the IBD patients there does not appear to be a correlation with differential expression of the EGR1 inhibitor NAB2. There also appears to be an aberrant stimulation response of *EGR1* expression to various inflammatory mediators in the IBD patients, in particular the CD patients. This study has demonstrated the localisation of EGR1 in both the cytoplasm and the nucleus in CRC cell lines and shown that EGR1 interacts with the CD susceptibility gene product NOD2 in SW480 and HRT18 cells, in the presence and absence of LPS. EGR1 has also been shown to interact with components of the cytoskeleton, α- and γ-tubulin, in SW480 cells. In this thesis we have found 22 potential novel interacting proteins of EGR1 using a yeast-two hybrid screen using a SW480 cell line library, and verified the interaction for two of these proteins, NPM1 and eEF1A1, in CRC cells. This yeast-two hybrid screen has identified several interesting novel proteins for further study in relation to EGR1 and CRC with many more interesting candidates to investigate which may help to identify new genes and novel pathways involved in CRC in order to further our understanding of the disease to aid the development of better diagnosis and new treatments. It is clear from our study that EGR1 may play a role in both CRC and

IBD, hence with a better understanding of the role that EGR1 plays and the pathways it regulates, it will hopefully lead to a greater understanding of colorectal diseases and new measures that can be taken to combat them.

7 References

Abdel-Latif, M. M., H. J. Windle, et al. (2004). "Helicobacter pylori activates the early growth response 1 protein in gastric epithelial cells." Infect Immun **72**(6): 3549-60.

Abdulkadir, S. A., J. M. Carbone, et al. (2001). "Frequent and early loss of the EGR1 corepressor NAB2 in human prostate carcinoma." Hum Pathol **32**(9): 935-9.

Abraham, C. and J. H. Cho (2006). "Functional consequences of NOD2 (CARD15) mutations." Inflamm Bowel Dis **12**(7): 641-50.

Adamson, E. D. and D. Mercola (2002). "Egr1 transcription factor: multiple roles in prostate tumor cell growth and survival." Tumour Biol **23**(2): 93-102.

Adamson, E. D., J. Yu, et al. (2005). "Co-factors p300 and CBP catch Egr1 in their network." Prostate **63**(4): 407-10.

Alvarez, A. and P. J. Woolf (2011). "RegNetB: predicting relevant regulator-gene relationships in localized prostate tumor samples." BMC Bioinformatics **12**: 243.

Amre, D. K., S. D'Souza, et al. (2007). "Imbalances in dietary consumption of fatty acids, vegetables, and fruits are associated with risk for Crohn's disease in children." Am J Gastroenterol **102**(9): 2016-25.

Andersen, V., A. Ernst, et al. (2010). "The polymorphism rs3024505 proximal to IL-10 is associated with risk of ulcerative colitis and Crohns disease in a Danish case-control study." BMC Med Genet **11**: 82.

Anderson, C. A., G. Boucher, et al. (2011). "Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47." Nat Genet **43**(3): 246-52.

Andersson, R. E., G. Olaison, et al. (2003). "Appendectomy is followed by increased risk of Crohn's disease." Gastroenterology **124**(1): 40-6.

Annese, V., M. R. Valvano, et al. (2006). "Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis." World J Gastroenterol **12**(23): 3636-44.

Aoki, K. and M. M. Taketo (2007). "Adenomatous polyposis coli (APC): a multi-functional tumor suppressor gene." J Cell Sci **120**(Pt 19): 3327-35.

Aramburu, J., K. Drews-Elger, et al. (2006). "Regulation of the hypertonic stress response and other cellular functions by the Rel-like transcription factor NFAT5." Biochem Pharmacol **72**(11): 1597-604.

Aust, D. E., J. P. Terdiman, et al. (2002). "The APC/beta-catenin pathway in ulcerative colitis-related colorectal carcinomas: a mutational analysis." Cancer **94**(5): 1421-7.

Bae, M. H., C. H. Jeong, et al. (2002). "Regulation of Egr-1 by association with the proteasome component C8." Biochim Biophys Acta **1592**(2): 163-7.

Baker, S. J., A. C. Preisinger, et al. (1990). "p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis." Cancer Res **50**(23): 7717-22.

Baldus, S. E., K. L. Schaefer, et al. (2010). "Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases." Clin Cancer Res **16**(3): 790-9.

Balk, S. P. and K. E. Knudsen (2008). "AR, the cell cycle, and prostate cancer." Nucl Recept Signal **6**: e001.
Barnich, N., J. E. Aguirre, et al. (2005). "Membrane recruitment of NOD2 in intestinal epithelial cells is essential for nuclear factor-{kappa}B activation in muramyl dipeptide recognition." J Cell Biol **170**(1): 21-6.
Baron, V., E. D. Adamson, et al. (2006). "The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGFbeta1, PTEN, p53, and fibronectin." Cancer Gene Ther **13**(2): 115-24.
Barrett, J. C., S. Hansoul, et al. (2008). "Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease." Nat Genet **40**(8): 955-62.
Bassi, M. T., M. P. Sperandeo, et al. (1999). "SLC7A8, a gene mapping within the lysinuric protein intolerance critical region, encodes a new member of the glycoprotein-associated amino acid transporter family." Genomics **62**(2): 297-303.
Baumgart, D. C., S. Thomas, et al. (2009). "Exaggerated inflammatory response of primary human myeloid dendritic cells to lipopolysaccharide in patients with inflammatory bowel disease." Clin Exp Immunol **157**(3): 423-36.
Beliakoff, J. and Z. Sun (2006). "Zimp7 and Zimp10, two novel PIAS-like proteins, function as androgen receptor coregulators." Nucl Recept Signal **4**: e017.
Benlloch, S., A. Paya, et al. (2006). "Detection of BRAF V600E mutation in colorectal cancer: comparison of automatic sequencing and real-time chemistry methodology." J Mol Diagn **8**(5): 540-3.
Bernstein, C. N., J. F. Blanchard, et al. (2001). "Cancer risk in patients with inflammatory bowel disease: a population-based study." Cancer **91**(4): 854-62.
Bernstein, C. N., P. Rawsthorne, et al. (2006). "A population-based case control study of potential risk factors for IBD." Am J Gastroenterol **101**(5): 993-1002.
Bertucci, F., S. Salas, et al. (2004). "Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters." Oncogene **23**(7): 1377-91.
Boland, C. R., S. N. Thibodeau, et al. (1998). "A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer." Cancer Res **58**(22): 5248-57.
Boles, K. S., R. Barten, et al. (1999). "Cloning of a new lectin-like receptor expressed on human NK cells." Immunogenetics **50**(1-2): 1-7.
Bosani, M., S. Ardizzone, et al. (2009). "Biologic targeting in the treatment of inflammatory bowel diseases." Biologics **3**: 77-97.
Bouma, G. and W. Strober (2003). "The immunological and genetic basis of inflammatory bowel disease." Nat Rev Immunol **3**(7): 521-33.
Boyko, E. J., M. K. Theis, et al. (1994). "Increased risk of inflammatory bowel disease associated with oral contraceptive use." Am J Epidemiol **140**(3): 268-78.
Brattain, M. G., W. D. Fine, et al. (1981). "Heterogeneity of malignant cells from a human colonic carcinoma." Cancer Res **41**(5): 1751-6.
Budarf, M. L., C. Labbe, et al. (2009). "GWA studies: rewriting the story of IBD." Trends Genet **25**(3): 137-46.

Calogero, A., A. Arcella, et al. (2001). "The early growth response gene EGR-1 behaves as a suppressor gene that is down-regulated independent of ARF/Mdm2 but not p53 alterations in fresh human gliomas." Clin Cancer Res **7**(9): 2788-96.

Cancer Research UK. (2011). "Bowel (colorectal) cancer - UK." Retrieved 14/06/2011, 2011, from http://info.cancerresearchuk.org/cancerstats/types/bowel/incidence/#source6.

Cao, X., G. Wei, et al. (2004). "The inositol 3-phosphatase PTEN negatively regulates Fc gamma receptor signaling, but supports Toll-like receptor 4 signaling in murine peritoneal macrophages." J Immunol **172**(8): 4851-7.

Cao, X. M., G. R. Guy, et al. (1992). "Regulation of the Egr-1 gene by tumor necrosis factor and interferons in primary human fibroblasts." J Biol Chem **267**(2): 1345-9.

Carpenter, B., K. J. Hill, et al. (2004). "BASP1 is a transcriptional cosuppressor for the Wilms' tumor suppressor protein WT1." Mol Cell Biol **24**(2): 537-49.

Caso, G., C. Barry, et al. (2009). "Dysregulation of CXCL9 and reduced tumor growth in Egr-1 deficient mice." J Hematol Oncol **2**: 7.

Chalhoub, N. and S. J. Baker (2009). "PTEN and the PI3-kinase pathway in cancer." Annu Rev Pathol **4**: 127-50.

Chang, W. C., R. A. Coudry, et al. (2007). "Loss of p53 enhances the induction of colitis-associated neoplasia by dextran sulfate sodium." Carcinogenesis **28**(11): 2375-81.

Chaudhuri, S., K. Vyas, et al. (2007). "Human ribosomal protein L13a is dispensable for canonical ribosome function but indispensable for efficient rRNA methylation." RNA **13**(12): 2224-37.

Chen, A., J. Xu, et al. (2006). "Curcumin inhibits human colon cancer cell growth by suppressing gene expression of epidermal growth factor receptor through reducing the activity of the transcription factor Egr-1." Oncogene **25**(2): 278-87.

Chen, M., S. K. Sastry, et al. (2011). "Src kinase pathway is involved in NFAT5-mediated S100A4 induction by hyperosmotic stress in colon cancer cells." Am J Physiol Cell Physiol **300**(5): C1155-63.

Choi, S., C. R. Jung, et al. (2008). "PRMT3 inhibits ubiquitination of ribosomal protein S2 and together forms an active enzyme complex." Biochim Biophys Acta **1780**(9): 1062-9.

Christy, B. and D. Nathans (1989). "DNA binding site of the growth factor-inducible protein Zif268." Proc Natl Acad Sci U S A **86**(22): 8737-41.

Christy, B. A., L. F. Lau, et al. (1988). "A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences." Proc Natl Acad Sci U S A **85**(21): 7857-61.

Chung, W., B. Kwabi-Addo, et al. (2008). "Identification of novel tumor markers in prostate, colon and breast cancer by unbiased methylation profiling." PLoS ONE **3**(4): e2079.

Clinton, M., L. Graeve, et al. (1991). "Evidence for nuclear targeting of prothymosin and parathymosin synthesized in situ." Proc Natl Acad Sci U S A **88**(15): 6608-12.

Clontech Laboratories Inc (2008) "pGBKT7 Vector Information." Protocol No. PT3248-5 **Volume**, DOI:

Clontech Laboratories Inc (2009) "pGADT7 Vector Information." Protocol No. PT3249-5 **Volume**, DOI:
Coleman, D. L., A. H. Bartiss, et al. (1992). "Lipopolysaccharide induces Egr-1 mRNA and protein in murine peritoneal macrophages." J Immunol **149**(9): 3045-51.
Collet, J. F., V. Stroobant, et al. (1999). "Mechanistic studies of phosphoserine phosphatase, an enzyme related to P-type ATPases." J Biol Chem **274**(48): 33985-90.
Collins, S., L. A. Wolfraim, et al. (2006). "Cutting Edge: TCR-induced NAB2 enhances T cell function by coactivating IL-2 transcription." J Immunol **177**(12): 8301-5.
Condeelis, J. (1995). "Elongation factor 1 alpha, translation and the cytoskeleton." Trends Biochem Sci **20**(5): 169-70.
Cosnes, J. (2004). "Tobacco and IBD: relevance in the understanding of disease mechanisms and clinical practice." Best Pract Res Clin Gastroenterol **18**(3): 481-96.
Cussenot, O., P. Berthon, et al. (1991). "Immortalization of human adult normal prostatic epithelial cells by liposomes containing large T-SV40 gene." J Urol **146**(3): 881-6.
Danese, S. and A. Mantovani (2010). "Inflammatory bowel disease and intestinal cancer: a paradigm of the Yin-Yang interplay between inflammation and cancer." Oncogene **29**(23): 3313-23.
Davies, H., G. R. Bignell, et al. (2002). "Mutations of the BRAF gene in human cancer." Nature **417**(6892): 949-54.
de Mestre, A. M., S. Rao, et al. (2005). "Early growth response gene 1 (EGR1) regulates heparanase gene transcription in tumor cells." J Biol Chem **280**(42): 35136-47.
de Vogel, S., V. Dindore, et al. (2008). "Dietary folate, methionine, riboflavin, and vitamin B-6 and risk of sporadic colorectal cancer." J Nutr **138**(12): 2372-8.
Deckmann, K., F. Rorsch, et al. (2010). "Dimethylcelecoxib inhibits mPGES-1 promoter activity by influencing EGR1 and NF-kappaB." Biochem Pharmacol **80**(9): 1365-72.
DeFranco, C., D. H. Damon, et al. (1993). "Nerve growth factor induces transcription of NGFIA through complex regulatory elements that are also sensitive to serum and phorbol 12-myristate 13-acetate." Mol Endocrinol **7**(3): 365-79.
Demeny, M. A., E. Soutoglou, et al. (2007). "Identification of a small TAF complex and its role in the assembly of TAF-containing complexes." PLoS One **2**(3): e316.
Dicuonzo, G., S. Angeletti, et al. (2001). "Colorectal carcinomas and PTEN/MMAC1 gene mutations." Clin Cancer Res **7**(12): 4049-53.
Din, F. V., M. G. Dunlop, et al. (2004). "Evidence for colorectal cancer cell specificity of aspirin effects on NF kappa B signalling and apoptosis." Br J Cancer **91**(2): 381-8.
Din, F. V., E. Theodoratou, et al. (2010). "Effect of aspirin and NSAIDs on risk and survival from colorectal cancer." Gut **59**(12): 1670-9.
Domingo, E., E. Espin, et al. (2004). "Activated BRAF targets proximal colon tumors with mismatch repair deficiency and MLH1 inactivation." Genes Chromosomes Cancer **39**(2): 138-42.

Drewinko, B., M. M. Romsdahl, et al. (1976). "Establishment of a human carcinoembryonic antigen-producing colon adenocarcinoma cell line." Cancer Res **36**(2 Pt 1): 467-75.
Du, W., W. Y. Li, et al. (2010). "Folate and fiber in the prevention of colorectal cancer: between shadows and the light." World J Gastroenterol **16**(8): 921-6.
Eaden, J. A., K. R. Abrams, et al. (2001). "The risk of colorectal cancer in ulcerative colitis: a meta-analysis." Gut **48**(4): 526-35.
Ejiri, S. (2002). "Moonlighting functions of polypeptide elongation factor 1: from actin bundling to zinc finger protein R1-associated nuclear localization." Biosci Biotechnol Biochem **66**(1): 1-21.
Ekbom, A., C. Helmick, et al. (1990). "Ulcerative colitis and colorectal cancer. A population-based study." N Engl J Med **323**(18): 1228-33.
Enokida, H., H. Shiina, et al. (2004). "CpG hypermethylation of MDR1 gene contributes to the pathogenesis and progression of human prostate cancer." Cancer Res **64**(17): 5956-62.
Esteller, M. (2007). "Cancer epigenomics: DNA methylomes and histone-modification maps." Nat Rev Genet **8**(4): 286-98.
Falini, B., C. Mecucci, et al. (2005). "Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype." N Engl J Med **352**(3): 254-66.
Fang, M., S. A. Wee, et al. (2007). "Evidence of EGR1 as a differentially expressed gene among proliferative skin diseases." Genomic Med **1**(1-2): 75-85.
Faour, W. H., N. Alaaeddine, et al. (2005). "Early growth response factor-1 mediates prostaglandin E2-dependent transcriptional suppression of cytokine-induced tumor necrosis factor-alpha gene expression in human macrophages and rheumatoid arthritis-affected synovial fibroblasts." J Biol Chem **280**(10): 9536-46.
Fearon, E. R. (2011). "Molecular genetics of colorectal cancer." Annu Rev Pathol **6**: 479-507.
Fearon, E. R. and B. Vogelstein (1990). "A genetic model for colorectal tumorigenesis." Cell **61**(5): 759-67.
Feinberg, A. P. and B. Tycko (2004). "The history of cancer epigenetics." Nat Rev Cancer **4**(2): 143-53.
Fiocchi, C. (1998). "Inflammatory bowel disease: etiology and pathogenesis." Gastroenterology **115**(1): 182-205.
Fodde, R., J. Kuipers, et al. (2001). "Mutations in the APC tumour suppressor gene cause chromosomal instability." Nat Cell Biol **3**(4): 433-8.
Fogh, J., J. M. Fogh, et al. (1977). "One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice." J Natl Cancer Inst **59**(1): 221-6.
Franke, A., D. P. McGovern, et al. (2011). "Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci." Nat Genet **42**(12): 1118-25.
Fritz, J. H., R. L. Ferrero, et al. (2006). "Nod-like proteins in immunity, inflammation and disease." Nat Immunol **7**(12): 1250-7.
Fu, M., X. Zhu, et al. (2003). "Egr-1 target genes in human endothelial cells identified by microarray analysis." Gene **315**: 33-41.

Galiatsatos, P. and W. D. Foulkes (2006). "Familial adenomatous polyposis." Am J Gastroenterol **101**(2): 385-98.
Gao, N., K. Asamitsu, et al. (2008). "AKIP1 enhances NF-kappaB-dependent gene expression by promoting the nuclear retention and phosphorylation of p65." J Biol Chem **283**(12): 7834-43.
Gashler, A. L., S. Swaminathan, et al. (1993). "A novel repression module, an extensive activation domain, and a bipartite nuclear localization signal defined in the immediate-early transcription factor Egr-1." Mol Cell Biol **13**(8): 4556-71.
Gaya, D. R., R. K. Russell, et al. (2006). "New genes in inflammatory bowel disease: lessons for complex diseases?" Lancet **367**(9518): 1271-84.
Geerling, B. J., P. C. Dagnelie, et al. (2000). "Diet as a risk factor for the development of ulcerative colitis." Am J Gastroenterol **95**(4): 1008-13.
Gent, A. E., M. D. Hellier, et al. (1994). "Inflammatory bowel disease and domestic hygiene in infancy." Lancet **343**(8900): 766-7.
Germain, C., F. Bihl, et al. (2010). "Characterization of alternatively spliced transcript variants of CLEC2D gene." J Biol Chem **285**(46): 36207-15.
Gjerset, R. A. (2006). "DNA damage, p14ARF, nucleophosmin (NPM/B23), and cancer." J Mol Histol **37**(5-7): 239-51.
Globocan. (2008). "Colorectal Cancer Incidence and Mortality Worldwide in 2008 " Retrieved 14/06/2011, from http://globocan.iarc.fr/factsheets/cancers/colorectal.asp.
Godet, P. G., G. R. May, et al. (1995). "Meta-analysis of the role of oral contraceptive agents in inflammatory bowel disease." Gut **37**(5): 668-73.
Gopalkrishnan, R. V., Z. Z. Su, et al. (1999). "Translational infidelity and human cancer: role of the PTI-1 oncogene." Int J Biochem Cell Biol **31**(1): 151-62.
Gordon, D., C. Abajian, et al. (1998). "Consed: a graphical tool for sequence finishing." Genome Res **8**(3): 195-202.
Grady, W. M. and J. M. Carethers (2008). "Genomic and epigenetic instability in colorectal cancer pathogenesis." Gastroenterology **135**(4): 1079-99.
Grady, W. M. and S. D. Markowitz (2002). "Genetic and epigenetic alterations in colon cancer." Annu Rev Genomics Hum Genet **3**: 101-28.
Granet, C. and P. Miossec (2004). "Combination of the pro-inflammatory cytokines IL-1, TNF-alpha and IL-17 leads to enhanced expression and additional recruitment of AP-1 family members, Egr-1 and NF-kappaB in osteoblast-like cells." Cytokine **26**(4): 169-77.
Green, L. M., K. J. Wagner, et al. (2009). "Dynamic interaction between WT1 and BASP1 in transcriptional regulation during differentiation." Nucleic Acids Res **37**(2): 431-40.
Gregg, J. L., K. E. Brown, et al. (2010). "Analysis of gene expression in prostate cancer epithelial and interstitial stromal cells using laser capture microdissection." BMC Cancer **10**: 165.
Grisendi, S., R. Bernardi, et al. (2005). "Role of nucleophosmin in embryonic development and tumorigenesis." Nature **437**(7055): 147-53.
Grisendi, S., C. Mecucci, et al. (2006). "Nucleophosmin and cancer." Nat Rev Cancer **6**(7): 493-505.
Groden, J., A. Thliveris, et al. (1991). "Identification and characterization of the familial adenomatous polyposis coli gene." Cell **66**(3): 589-600.

Guanti, G., N. Resta, et al. (2000). "Involvement of PTEN mutations in the genetic pathways of colorectal cancerogenesis." Hum Mol Genet **9**(2): 283-7.
Guermah, M., K. Ge, et al. (2003). "The TBN protein, which is essential for early embryonic mouse development, is an inducible TAFII implicated in adipogenesis." Mol Cell **12**(4): 991-1001.
Guha, M., M. A. O'Connell, et al. (2001). "Lipopolysaccharide activation of the MEK-ERK1/2 pathway in human monocytic cells mediates tissue factor and tumor necrosis factor alpha expression by inducing Elk-1 phosphorylation and Egr-1 expression." Blood **98**(5): 1429-39.
Gupta, K. K., S. S. Bharne, et al. (2006). "Dietary antioxidant curcumin inhibits microtubule assembly through tubulin binding." FEBS J **273**(23): 5320-32.
Habermann, J. K., U. Paulsen, et al. (2007). "Stage-specific alterations of the genome, transcriptome, and proteome during colorectal carcinogenesis." Genes Chromosomes Cancer **46**(1): 10-26.
Halfvarson, J., L. Bodin, et al. (2003). "Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics." Gastroenterology **124**(7): 1767-73.
Ham, J., A. Eilers, et al. (2000). "c-Jun and the transcriptional control of neuronal apoptosis." Biochem Pharmacol **60**(8): 1015-21.
Hampe, J., K. Heymann, et al. (2003). "Association of inflammatory bowel disease with indicators for childhood antigen and infection exposure." Int J Colorectal Dis **18**(5): 413-7.
Hartl, M., A. Nist, et al. (2009). "Inhibition of Myc-induced cell transformation by brain acid-soluble protein 1 (BASP1)." Proc Natl Acad Sci U S A **106**(14): 5604-9.
Heyer, J., K. Yang, et al. (1999). "Mouse models for colorectal cancer." Oncogene **18**(38): 5325-33.
Holzmann, K., B. Klump, et al. (1998). "Comparative analysis of histology, DNA content, p53 and Ki-ras mutations in colectomy specimens with long-standing ulcerative colitis." Int J Cancer **76**(1): 1-6.
Hong, Y., T. Downey, et al. (2010). "A 'metastasis-prone' signature for early-stage mismatch-repair proficient sporadic colorectal cancer patients and its implications for possible therapeutics." Clin Exp Metastasis **27**(2): 83-90.
Hong, Y., K. S. Ho, et al. (2007). "A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: implication for tumorigenesis." Clin Cancer Res **13**(4): 1107-14.
Hothorn, T., F. Bretz, et al. (2008). "Simultaneous inference in general parametric models." Biom J **50**(3): 346-63.
Hotokezaka, Y., K. van Leyen, et al. (2009). "alphaNAC depletion as an initiator of ER stress-induced apoptosis in hypoxia." Cell Death Differ **16**(11): 1505-14.
Houlston, R. S., J. Cheadle, et al. (2010). "Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33." Nat Genet **42**(11): 973-7.
Houston, P., C. J. Campbell, et al. (2001). "The transcriptional corepressor NAB2 blocks Egr-1-mediated growth factor activation and angiogenesis." Biochem Biophys Res Commun **283**(2): 480-6.

Huang, C. Y., J. Beliakoff, et al. (2005). "hZimp7, a novel PIAS-like protein, enhances androgen receptor-mediated transcription and interacts with SWI/SNF-like BAF complexes." Mol Endocrinol **19**(12): 2915-29.
Huang da, W., B. T. Sherman, et al. (2009). "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." Nucleic Acids Res **37**(1): 1-13.
Huang da, W., B. T. Sherman, et al. (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc **4**(1): 44-57.
Huang, R. P., Y. Fan, et al. (1997). "Decreased Egr-1 expression in human, mouse and rat mammary cells and tissues correlates with tumor formation." Int J Cancer **72**(1): 102-9.
Huang, R. P., C. Liu, et al. (1995). "Egr-1 negatively regulates human tumor cell growth via the DNA-binding domain." Cancer Res **55**(21): 5054-62.
Hussain, S. P., P. Amstad, et al. (2000). "Increased p53 mutation load in noncancerous colon tissue from ulcerative colitis: a cancer-prone chronic inflammatory disease." Cancer Res **60**(13): 3333-7.
Huxley, R. R., A. Ansary-Moghaddam, et al. (2009). "The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence." Int J Cancer **125**(1): 171-80.
Ilyas, M., J. Straub, et al. (1999). "Genetic pathways in colorectal and other cancers." Eur J Cancer **35**(14): 1986-2002.
Ishihara, S., M. M. Aziz, et al. (2009). "Inflammatory bowel disease: review from the aspect of genetics." J Gastroenterol **44**(11): 1097-108.
Issa, J. P. (2008). "Colon cancer: it's CIN or CIMP." Clin Cancer Res **14**(19): 5939-40.
Jass, J. R. (2007). "Classification of colorectal cancer based on correlation of clinical, morphological and molecular features." Histopathology **50**(1): 113-30.
Jauliac, S., C. Lopez-Rodriguez, et al. (2002). "The role of NFAT transcription factors in integrin-mediated carcinoma invasion." Nat Cell Biol **4**(7): 540-4.
Jess, T., M. Gamborg, et al. (2005). "Increased risk of intestinal cancer in Crohn's disease: a meta-analysis of population-based cohort studies." Am J Gastroenterol **100**(12): 2724-9.
Ji, X., J. Tang, et al. (2010). "Distinguishing between cancer driver and passenger gene alteration candidates via cross-species comparison: a pilot study." BMC Cancer **10**: 426.
Jin, C., Y. Zhang, et al. (2005). "Human Yip1A specifies the localization of Yif1 to the Golgi apparatus." Biochem Biophys Res Commun **334**(1): 16-22.
Jo, D., M. S. Lyu, et al. (2001). "Identification and genetic mapping of the mouse Fkbp9 gene encoding a new member of FK506-binding protein family." Mol Cells **12**(2): 272-5.
Jones, P. A. and S. B. Baylin (2002). "The fundamental role of epigenetic events in cancer." Nat Rev Genet **3**(6): 415-28.
Kadl, A., J. Huber, et al. (2002). "Analysis of inflammatory gene induction by oxidized phospholipids in vivo by quantitative real-time RT-PCR in comparison with effects of LPS." Vascul Pharmacol **38**(4): 219-27.

Kaighn, M. E., K. S. Narayan, et al. (1979). "Establishment and characterization of a human prostatic carcinoma cell line (PC-3)." Invest Urol **17**(1): 16-23.
Kakar, S., G. Deng, et al. (2008). "Clinicopathologic characteristics, CpG island methylator phenotype, and BRAF mutations in microsatellite-stable colorectal cancers without chromosomal instability." Arch Pathol Lab Med **132**(6): 958-64.
Kambara, T., L. A. Simms, et al. (2004). "BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum." Gut **53**(8): 1137-44.
Kang, C. B., Y. Hong, et al. (2008). "FKBP family proteins: immunophilins with versatile biological functions." Neurosignals **16**(4): 318-25.
Kaplan, K. B., A. A. Burds, et al. (2001). "A role for the Adenomatous Polyposis Coli protein in chromosome segregation." Nat Cell Biol **3**(4): 429-32.
Kapp, L. D. and J. R. Lorsch (2004). "The molecular mechanics of eukaryotic translation." Annu Rev Biochem **73**: 657-704.
Kern, S. E., M. Redston, et al. (1994). "Molecular genetic profiles of colitis-associated neoplasms." Gastroenterology **107**(2): 420-8.
Kerr, L. E., J. L. Birse-Archbold, et al. (2007). "Nucleophosmin is a novel Bax chaperone that regulates apoptotic cell death." Oncogene **26**(18): 2554-62.
Key, T. J. (2011). "Fruit and vegetables and cancer risk." Br J Cancer **104**(1): 6-11.
Kim, M. S., J. Lee, et al. (2010). "DNA methylation markers in colorectal cancer." Cancer Metastasis Rev **29**(1): 181-206.
Kinzler, K. W. and B. Vogelstein (1996). "Lessons from hereditary colorectal cancer." Cell **87**(2): 159-70.
Kirsch, K. H., Y. Korradi, et al. (1996). "Mader: a novel nuclear protein over expressed in human melanomas." Oncogene **12**(5): 963-71.
Kitching, R., H. Li, et al. (2003). "Characterization of a novel human breast cancer associated gene (BCA3) encoding an alternatively spliced proline-rich protein." Biochim Biophys Acta **1625**(1): 116-21.
Kobayashi, D., M. Yamada, et al. (2002). "Overexpression of early growth response-1 as a metastasis-regulatory factor in gastric cancer." Anticancer Res **22**(6C): 3963-70.
Kondili, K., O. Tsolas, et al. (1996). "Selective interaction between parathymosin and histone H1." Eur J Biochem **242**(1): 67-74.
Kraus, S. and N. Arber (2009). "Inflammation and colorectal cancer." Curr Opin Pharmacol **9**(4): 405-10.
Krechowec, S. and A. Plagge (2008). "Physiological dysfunctions associated with mutations of the imprinted Gnas locus." Physiology (Bethesda) **23**: 221-9.
Krones-Herzig, A., S. Mittal, et al. (2005). "Early growth response 1 acts as a tumor suppressor in vivo and in vitro via regulation of p53." Cancer Res **65**(12): 5133-43.
Kumbrink, J., M. Gerlinger, et al. (2005). "Egr-1 induces the expression of its corepressor nab2 by activation of the nab2 promoter thereby establishing a negative feedback loop." J Biol Chem **280**(52): 42785-93.
Kumbrink, J., K. H. Kirsch, et al. (2010). "EGR1, EGR2, and EGR3 activate the expression of their coregulator NAB2 establishing a negative feedback loop in cells of neuroectodermal and epithelial origin." J Cell Biochem **111**(1): 207-17.

Kuriyama, R., M. Bettencourt-Dias, et al. (2009). "Gamma-tubulin-containing abnormal centrioles are induced by insufficient Plk4 in human HCT116 colorectal cancer cells." J Cell Sci **122**(Pt 12): 2014-23.
Kurzawski, G., J. Suchy, et al. (2004). "The NOD2 3020insC mutation and the risk of colorectal cancer." Cancer Res **64**(5): 1604-6.
Lamberti, A., M. Caraglia, et al. (2004). "The translation elongation factor 1A in tumorigenesis, signal transduction and apoptosis: review article." Amino Acids **26**(4): 443-8.
Larsson, S. C. and A. Wolk (2006). "Meat consumption and risk of colorectal cancer: a meta-analysis of prospective studies." Int J Cancer **119**(11): 2657-64.
Lascorz, J., A. Forsti, et al. (2010). "Genome-wide association study for colorectal cancer identifies risk polymorphisms in German familial cases and implicates MAPK signalling pathways in disease susceptibility." Carcinogenesis **31**(9): 1612-9.
Lea, I. A., M. A. Jackson, et al. (2009). "Genetic pathways to colorectal cancer." Mutat Res **670**(1-2): 96-8.
Lee, S., N. Y. Cho, et al. (2008). "Clinicopathological features of CpG island methylator phenotype-positive colorectal cancer and its adverse prognosis in relation to KRAS/BRAF mutation." Pathol Int **58**(2): 104-13.
Lee, S., A. M. Francoeur, et al. (1992). "Tissue-specific expression in mammalian brain, heart, and muscle of S1, a member of the elongation factor-1 alpha gene family." J Biol Chem **267**(33): 24064-8.
Lee, S. L., L. C. Tourtellotte, et al. (1995). "Growth and differentiation proceeds normally in cells deficient in the immediate early gene NGFI-A." J Biol Chem **270**(17): 9971-7.
Legrand-Poels, S., G. Kustermans, et al. (2007). "Modulation of Nod2-dependent NF-kappaB signaling by the actin cytoskeleton." J Cell Sci **120**(Pt 7): 1299-310.
Leibovitz, A., J. C. Stinson, et al. (1976). "Classification of human colorectal adenocarcinoma cell lines." Cancer Res **36**(12): 4562-9.
Lengauer, C., K. W. Kinzler, et al. (1998). "Genetic instabilities in human cancers." Nature **396**(6712): 643-9.
Leon, J., J. Casado, et al. (2011). "Gender-related invasion differences associated with mRNA expression levels of melatonin membrane receptors in colorectal cancer." Mol Carcinog.
Leung, E., J. Hong, et al. (2007). "Splicing of NOD2 (CARD15) RNA transcripts." Mol Immunol **44**(4): 284-94.
Li, L. C. and R. Dahiya (2002). "MethPrimer: designing primers for methylation PCRs." Bioinformatics **18**(11): 1427-31.
Liang, P. S., T. Y. Chen, et al. (2009). "Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis." Int J Cancer **124**(10): 2406-15.
Liu, C., J. Yao, et al. (1999). "The transcription factor EGR-1 suppresses transformation of human fibrosarcoma HT1080 cells by coordinated induction of transforming growth factor-beta1, fibronectin, and plasminogen activator inhibitor-1." J Biol Chem **274**(7): 4400-11.

Loftus, E. V., Jr. (2004). "Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences." Gastroenterology **126**(6): 1504-17.
Lopez, S., L. Stuhl, et al. (2005). "NACA is a positive regulator of human erythroid-cell differentiation." J Cell Sci **118**(Pt 8): 1595-605.
Lopez-Rodriguez, C., J. Aramburu, et al. (2001). "Bridging the NFAT and NF-kappaB families: NFAT5 dimerization regulates cytokine gene transcription in response to osmotic stress." Immunity **15**(1): 47-58.
Lu, Y., T. Li, et al. (2011). "Early growth response 1 (Egr-1) regulates phosphorylation of microtubule-associated protein tau in mammalian brain." J Biol Chem **286**(23): 20569-81.
Luther, J., M. Dave, et al. (2010). "Association between Helicobacter pylori infection and inflammatory bowel disease: a meta-analysis and systematic review of the literature." Inflamm Bowel Dis **16**(6): 1077-84.
Luyendyk, J. P., G. A. Schabbauer, et al. (2008). "Genetic analysis of the role of the PI3K-Akt pathway in lipopolysaccharide-induced cytokine and tissue factor gene expression in monocytes/macrophages." J Immunol **180**(6): 4218-26.
Lynch, H. T. and A. de la Chapelle (2003). "Hereditary colorectal cancer." N Engl J Med **348**(10): 919-32.
Macian, F., C. Lopez-Rodriguez, et al. (2001). "Partners in transcription: NFAT and AP-1." Oncogene **20**(19): 2476-89.
Macias, E., A. Jin, et al. "An ARF-independent c-MYC-activated tumor suppression pathway mediated by ribosomal protein-Mdm2 Interaction." Cancer Cell **18**(3): 231-43.
Maeda, S., L. C. Hsu, et al. (2005). "Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing." Science **307**(5710): 734-8.
Maegawa, M., T. Arao, et al. (2009). "EGFR mutation up-regulates EGR1 expression through the ERK pathway." Anticancer Res **29**(4): 1111-7.
Mahalingam, D., A. Natoni, et al. (2010). "Early growth response-1 is a regulator of DR5-induced apoptosis in colon cancer cells." Br J Cancer **102**(4): 754-64.
Mahid, S. S., K. S. Minor, et al. (2006). "Smoking and inflammatory bowel disease: a meta-analysis." Mayo Clin Proc **81**(11): 1462-71.
Malumbres, M. and M. Barbacid (2003). "RAS oncogenes: the first 30 years." Nat Rev Cancer **3**(6): 459-65.
Marintchev, A., K. A. Edmonds, et al. (2009). "Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation." Cell **136**(3): 447-60.
Markowitz, S., J. Wang, et al. (1995). "Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability." Science **268**(5215): 1336-8.
Martic, G., Z. Karetsou, et al. (2005). "Parathymosin affects the binding of linker histone H1 to nucleosomes and remodels chromatin structure." J Biol Chem **280**(16): 16143-50.
Mathew, P. A., S. S. Chuang, et al. (2004). "The LLT1 receptor induces IFN-gamma production by human natural killer cells." Mol Immunol **40**(16): 1157-63.
Matsunoshita, Y., K. Ijiri, et al. (2011). "Suppression of osteosarcoma cell invasion by chemotherapy is mediated by urokinase plasminogen activator activity via up-regulation of EGR1." PLoS One **6**(1): e16234.

Mawdsley, J. E. and D. S. Rampton (2005). "Psychological stress in IBD: new insights into pathogenic and therapeutic implications." Gut **54**(10): 1481-91.
Mazumder, B., P. Sampath, et al. (2003). "Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control." Cell **115**(2): 187-98.
McBain, J. A., J. L. Weese, et al. (1984). "Establishment and characterization of human colorectal cancer cell lines." Cancer Res **44**(12 Pt 1): 5813-21.
McCarroll, S. A., A. Huett, et al. (2008). "Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease." Nat Genet **40**(9): 1107-12.
McCoy, C., D. E. Smith, et al. (1995). "12-O-tetradecanoylphorbol-13-acetate activation of the MDR1 promoter is mediated by EGR1." Mol Cell Biol **15**(11): 6100-8.
Migliore, L., F. Migheli, et al. (2011). "Genetics, cytogenetics, and epigenetics of colorectal cancer." J Biomed Biotechnol **2011**: 792362.
Miyoshi, Y., H. Nagase, et al. (1992). "Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene." Hum Mol Genet **1**(4): 229-33.
Mizoguchi, A. and E. Mizoguchi (2010). "Animal models of IBD: linkage to human disease." Curr Opin Pharmacol **10**(5): 578-87.
Moghaddam, A. A., M. Woodward, et al. (2007). "Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events." Cancer Epidemiol Biomarkers Prev **16**(12): 2533-47.
Molnar, G., A. Crozat, et al. (1994). "The immediate-early gene Egr-1 regulates the activity of the thymidine kinase promoter at the G0-to-G1 transition of the cell cycle." Mol Cell Biol **14**(8): 5242-8.
Molodecky, N. A. and G. G. Kaplan (2010). "Environmental risk factors for inflammatory bowel disease." Gastroenterol Hepatol (N Y) **6**(5): 339-46.
Mora, G. R., K. R. Olivier, et al. (2004). "The cytoskeleton differentially localizes the early growth response gene-1 protein in cancer and benign cells of the prostate." Mol Cancer Res **2**(2): 115-28.
Moribe, T., N. Iizuka, et al. (2008). "Identification of novel aberrant methylation of BASP1 and SRD5A2 for early diagnosis of hepatocellular carcinoma by genome-wide search." Int J Oncol **33**(5): 949-58.
Morin, P. J., B. Vogelstein, et al. (1996). "Apoptosis and APC in colorectal tumorigenesis." Proc Natl Acad Sci U S A **93**(15): 7950-4.
Muller, M. R. and A. Rao (2010). "NFAT, immunity and cancer: a transcription factor comes of age." Nat Rev Immunol **10**(9): 645-56.
Munkholm, P. (2003). "Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease." Aliment Pharmacol Ther **18 Suppl 2**: 1-5.
Nagasaka, T., H. Sasamoto, et al. (2004). "Colorectal cancer with mutation in BRAF, KRAS, and wild-type with respect to both oncogenes showing different patterns of DNA methylation." J Clin Oncol **22**(22): 4584-94.
Nakano, I., J. D. Dougherty, et al. (2007). "Phosphoserine phosphatase is expressed in the neural stem cell niche and regulates neural stem and progenitor cell proliferation." Stem Cells **25**(8): 1975-84.

Naraba, H., C. Yokoyama, et al. (2002). "Transcriptional regulation of the membrane-associated prostaglandin E2 synthase gene. Essential role of the transcription factor Egr-1." J Biol Chem **277**(32): 28601-8.
Nishi, H., K. H. Nishi, et al. (2002). "Early Growth Response-1 gene mediates up-regulation of epidermal growth factor receptor expression during hypoxia." Cancer Res **62**(3): 827-34.
Nishisho, I., Y. Nakamura, et al. (1991). "Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients." Science **253**(5020): 665-9.
Noble, C., E. Nimmo, et al. (2006). "Novel susceptibility genes in inflammatory bowel disease." World J Gastroenterol **12**(13): 1991-9.
Noguchi, E., Y. Homma, et al. (2009). "A Crohn's disease-associated NOD2 mutation suppresses transcription of human IL10 by inhibiting activity of the nuclear ribonucleoprotein hnRNP-A1." Nat Immunol **10**(5): 471-9.
Nozawa, Y., N. Van Belzen, et al. (1996). "Expression of nucleophosmin/B23 in normal and neoplastic colorectal mucosa." J Pathol **178**(1): 48-52.
Ogishima, T., H. Shiina, et al. (2005). "Increased heparanase expression is caused by promoter hypomethylation and up-regulation of transcriptional factor early growth response-1 in human prostate cancer." Clin Cancer Res **11**(3): 1028-36.
Ogishima, T., H. Shiina, et al. (2005). "Promoter CpG hypomethylation and transcription factor EGR1 hyperactivate heparanase expression in bladder cancer." Oncogene **24**(45): 6765-72.
Ogura, Y., D. K. Bonen, et al. (2001). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.
Pambuccian, C. A., G. M. Oprea, et al. (2002). "Reduced expression of early growth response-1 gene in leiomyoma as identified by mRNA differential display." Gynecol Oncol **84**(3): 431-6.
Pavletich, N. P. and C. O. Pabo (1991). "Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A." Science **252**(5007): 809-17.
Pawlinski, R., B. Pedersen, et al. (2003). "Regulation of tissue factor and inflammatory mediators by Egr-1 in a mouse endotoxemia model." Blood **101**(10): 3940-7.
Peeraer, Y., A. Rabijns, et al. (2002). "Purification, crystallization and preliminary X-ray diffraction analysis of human phosphoserine phosphatase." Acta Crystallogr D Biol Crystallogr **58**(Pt 1): 133-4.
Peng, Y., J. Lee, et al. (2010). "A novel role for protein inhibitor of activated STAT (PIAS) proteins in modulating the activity of Zimp7, a novel PIAS-like protein, in androgen receptor-mediated transcription." J Biol Chem **285**(15): 11465-75.
Pineda, M., E. Fernandez, et al. (1999). "Identification of a membrane protein, LAT-2, that Co-expresses with 4F2 heavy chain, an L-type amino acid transport activity with broad specificity for small and large zwitterionic amino acids." J Biol Chem **274**(28): 19738-44.
Pinheiro, J., D. Bates, et al. (2009) "nlme: Linear and Nonlinear Mixed Effects Models." R package version 3.1-96 **Volume**, DOI:
Pipaon, C., P. J. Real, et al. (2005). "Defective binding of transcriptional repressor ZEB via DNA methylation contributes to increased constitutive levels of p73 in Fanconi anemia cells." FEBS Lett **579**(21): 4610-4.

Polakis, P. (1995). "Mutations in the APC gene and their implications for protein structure and function." Curr Opin Genet Dev **5**(1): 66-71.
Potter, J. D. (1999). "Colorectal cancer: molecules and populations." J Natl Cancer Inst **91**(11): 916-32.
Powell, S. M., N. Zilz, et al. (1992). "APC mutations occur early during colorectal tumorigenesis." Nature **359**(6392): 235-7.
R Development Core Team. (2009). "R: A language and environment for statistical computing." from http://www.R-project.org.
Rhodes, D. R., J. Yu, et al. (2004). "ONCOMINE: a cancer microarray database and integrated data-mining platform." Neoplasia **6**(1): 1-6.
Riboli, E. and T. Norat (2003). "Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk." Am J Clin Nutr **78**(3 Suppl): 559S-569S.
Roda, G., A. Sartini, et al. (2010). "Intestinal epithelial cells in inflammatory bowel diseases." World J Gastroenterol **16**(34): 4264-71.
Rodriguez, J., J. Frigola, et al. (2006). "Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers." Cancer Res **66**(17): 8462-9468.
Rodriguez, S., O. Jafer, et al. (2003). "Expression profile of genes from 12p in testicular germ cell tumors of adolescents and adults associated with i(12p) and amplification at 12p11.2-p12.1." Oncogene **22**(12): 1880-91.
Rolli, M., A. Kotlyarov, et al. (1999). "Stress-induced stimulation of early growth response gene-1 by p38/stress-activated protein kinase 2 is mediated by a cAMP-responsive promoter element in a MAPKAP kinase 2-independent manner." J Biol Chem **274**(28): 19559-64.
Rosenstiel, P., K. Huse, et al. (2007). "Functional characterization of two novel 5' untranslated exons reveals a complex regulation of NOD2 protein expression." BMC Genomics **8**: 472.
Rossier, G., C. Meier, et al. (1999). "LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine." J Biol Chem **274**(49): 34948-54.
Rozovsky, N., A. C. Butterworth, et al. (2008). "Interactions between eIF4AI and its accessory factors eIF4B and eIF4H." RNA **14**(10): 2136-48.
Rubbi, C. P. and J. Milner (2003). "Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses." EMBO J **22**(22): 6068-77.
Russell, R. K. and J. Satsangi (2004). "IBD: a family affair." Best Pract Res Clin Gastroenterol **18**(3): 525-39.
Russo, M. W., B. R. Sevetson, et al. (1995). "Identification of NAB1, a repressor of NGFI-A- and Krox20-mediated transcription." Proc Natl Acad Sci U S A **92**(15): 6873-7.
Rutgeerts, P., G. D'Haens, et al. (1994). "Appendectomy protects against ulcerative colitis." Gastroenterology **106**(5): 1251-3.
Sabbah, A., T. H. Chang, et al. (2009). "Activation of innate immune antiviral responses by Nod2." Nat Immunol **10**(10): 1073-80.
Saegusa, M., M. Hashimura, et al. (2008). "Transcription factor Egr1 acts as an upstream regulator of beta-catenin signalling through up-regulation of TCF4 and p300 expression during trans-differentiation of endometrial carcinoma cells." J Pathol **216**(4): 521-32.

Samad, A. K., R. S. Taylor, et al. (2005). "A meta-analysis of the association of physical activity with reduced risk of colorectal cancer." Colorectal Dis **7**(3): 204-13.
Samuels, Y., Z. Wang, et al. (2004). "High frequency of mutations of the PIK3CA gene in human cancers." Science **304**(5670): 554.
Sanchez-Beato, M., E. Sanchez, et al. (2006). "Variability in the expression of polycomb proteins in different normal and tumoral tissues. A pilot study using tissue microarrays." Mod Pathol **19**(5): 684-94.
Sastri, M., D. M. Barraclough, et al. (2005). "A-kinase-interacting protein localizes protein kinase A in the nucleus." Proc Natl Acad Sci U S A **102**(2): 349-54.
Sauer, L., D. Gitenay, et al. (2010). "Mutant p53 initiates a feedback loop that involves Egr-1/EGF receptor/ERK in prostate cancer cells." Oncogene **29**(18): 2628-37.
Scharnhorst, V., A. L. Menke, et al. (2000). "EGR-1 enhances tumor growth and modulates the effect of the Wilms' tumor 1 gene products on tumorigenicity." Oncogene **19**(6): 791-800.
Schuebel, K. E., W. Chen, et al. (2007). "Comparing the DNA hypermethylome with gene mutations in human colorectal cancer." PLoS Genet **3**(9): 1709-23.
Scotto, L., G. Narayan, et al. (2008). "Integrative genomics analysis of chromosome 5p gain in cervical cancer reveals target over-expressed genes, including Drosha." Mol Cancer **7**: 58.
Segditsas, S. and I. Tomlinson (2006). "Colorectal cancer and genetic alterations in the Wnt pathway." Oncogene **25**(57): 7531-7.
Seth, R., S. Crook, et al. (2009). "Concomitant mutations and splice variants in KRAS and BRAF demonstrate complex perturbation of the Ras/Raf signalling pathway in advanced colorectal cancer." Gut **58**(9): 1234-41.
Seyfert, V. L., S. B. McMahon, et al. (1990). "Methylation of an immediate-early inducible gene as a mechanism for B cell tolerance induction." Science **250**(4982): 797-800.
Shi, L., R. Kishore, et al. (2002). "Lipopolysaccharide stimulation of ERK1/2 increases TNF-alpha production via Egr-1." Am J Physiol Cell Physiol **282**(6): C1205-11.
Shibata, D., M. A. Peinado, et al. (1994). "Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation." Nat Genet **6**(3): 273-81.
Simon, J. A. and R. E. Kingston (2009). "Mechanisms of polycomb gene silencing: knowns and unknowns." Nat Rev Mol Cell Biol **10**(10): 697-708.
Sjoblom, T., S. Jones, et al. (2006). "The consensus coding sequences of human breast and colorectal cancers." Science **314**(5797): 268-74.
Skrzypczak, M., K. Goryca, et al. (2010). "Modeling oncogenic signaling in colon tumors by multidirectional analyses of microarray data directed for maximization of analytical reliability." PLoS One **5**(10).
Smith, K. J., K. A. Johnson, et al. (1993). "The APC gene product in normal and tumor cells." Proc Natl Acad Sci U S A **90**(7): 2846-50.
Smith, K. J., D. B. Levy, et al. (1994). "Wild-type but not mutant APC associates with the microtubule cytoskeleton." Cancer Res **54**(14): 3672-5.

Soares, D. C., P. N. Barlow, et al. (2009). "Structural models of human eEF1A1 and eEF1A2 reveal two distinct surface clusters of sequence variation and potential differences in phosphorylation." PLoS One **4**(7): e6315.
Soule, H. D., J. Vazguez, et al. (1973). "A human cell line from a pleural effusion derived from a breast carcinoma." J Natl Cancer Inst **51**(5): 1409-16.
Soulieres, D., W. Greer, et al. (2010). "KRAS mutation testing in the treatment of metastatic colorectal cancer with anti-EGFR therapies." Curr Oncol **17 Suppl 1**: S31-40.
Soutoglou, E., M. A. Demeny, et al. (2005). "The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners." Mol Cell Biol **25**(10): 4092-104.
Srinivasan, R., G. M. Mager, et al. (2006). "NAB2 represses transcription by interacting with the CHD4 subunit of the nucleosome remodeling and deacetylase (NuRD) complex." J Biol Chem **281**(22): 15129-37.
Starr, T. K., R. Allaei, et al. (2009). "A transposon-based genetic screen in mice identifies genes altered in colorectal cancer." Science **323**(5922): 1747-50.
Stilo, R., D. Liguoro, et al. (2003). "The alpha-chain of the nascent polypeptide-associated complex binds to and regulates FADD function." Biochem Biophys Res Commun **303**(4): 1034-41.
Stoler, D. L., N. Chen, et al. (1999). "The onset and extent of genomic instability in sporadic colorectal tumor progression." Proc Natl Acad Sci U S A **96**(26): 15121-6.
Stone, K. R., D. D. Mickey, et al. (1978). "Isolation of a human prostate carcinoma cell line (DU 145)." Int J Cancer **21**(3): 274-81.
Strober, W., I. J. Fuss, et al. (2002). "The immunology of mucosal models of inflammation." Annu Rev Immunol **20**: 495-549.
Strober, W., P. J. Murray, et al. (2006). "Signalling pathways and molecular interactions of NOD1 and NOD2." Nat Rev Immunol **6**(1): 9-20.
Su, A. I., T. Wiltshire, et al. (2004). "A gene atlas of the mouse and human protein-encoding transcriptomes." Proc Natl Acad Sci U S A **101**(16): 6062-7.
Subbaramaiah, K., K. Yoshimatsu, et al. (2004). "Microsomal prostaglandin E synthase-1 is overexpressed in inflammatory bowel disease. Evidence for involvement of the transcription factor Egr-1." J Biol Chem **279**(13): 12647-58.
Sugimoto, T., N. Seki, et al. (2009). "The galanin signaling cascade is a candidate pathway regulating oncogenesis in human squamous cell carcinoma." Genes Chromosomes Cancer **48**(2): 132-42.
Sukhatme, V. P. (1990). "Early transcriptional events in cell growth: the Egr family." J Am Soc Nephrol **1**(6): 859-66.
Sukhatme, V. P., X. M. Cao, et al. (1988). "A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization." Cell **53**(1): 37-43.
Svaren, J., T. Ehrig, et al. (2000). "EGR1 target genes in prostate carcinoma cells identified by microarray analysis." J Biol Chem **275**(49): 38524-31.
Svaren, J., B. R. Sevetson, et al. (1996). "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by proliferative and differentiative stimuli." Mol Cell Biol **16**(7): 3545-53.

Svaren, J., B. R. Sevetson, et al. (1998). "Novel mutants of NAB corepressors enhance activation by Egr transactivators." EMBO J **17**(20): 6010-9.
Swirnoff, A. H., E. D. Apel, et al. (1998). "Nab1, a corepressor of NGFI-A (Egr-1), contains an active transcriptional repression domain." Mol Cell Biol **18**(1): 512-24.
Taketo, M. M. and W. Edelmann (2009). "Mouse models of colon cancer." Gastroenterology **136**(3): 780-98.
Tanabe, T., M. Chamaillard, et al. (2004). "Regulatory regions and critical residues of NOD2 involved in muramyl dipeptide recognition." EMBO J **23**(7): 1587-97.
ten Dijke, P. and C. S. Hill (2004). "New insights into TGF-beta-Smad signalling." Trends Biochem Sci **29**(5): 265-73.
Tenesa, A. and M. G. Dunlop (2009). "New insights into the aetiology of colorectal cancer from genome-wide association studies." Nat Rev Genet **10**(6): 353-8.
Terry, P., E. Giovannucci, et al. (2001). "Fruit, vegetables, dietary fiber, and risk of colorectal cancer." J Natl Cancer Inst **93**(7): 525-33.
The Wellcome Trust Case Control Consortium (2007). "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." Nature **447**(7145): 661-78.
Theodoratou, E., S. M. Farrington, et al. (2008). "Dietary vitamin B6 intake and the risk of colorectal cancer." Cancer Epidemiol Biomarkers Prev **17**(1): 171-82.
Theodoratou, E., S. M. Farrington, et al. (2008). "Modification of the inverse association between dietary vitamin D intake and colorectal cancer risk by a FokI variant supports a chemoprotective action of Vitamin D intake mediated through VDR binding." Int J Cancer **123**(9): 2170-9.
Theodoratou, E., G. McNeill, et al. (2007). "Dietary fatty acids and colorectal cancer: a case-control study." Am J Epidemiol **166**(2): 181-95.
Thiel, G. and G. Cibelli (2002). "Regulation of life and death by the zinc finger transcription factor Egr-1." J Cell Physiol **193**(3): 287-92.
Thompson, N. P., R. Driscoll, et al. (1996). "Genetics versus environment in inflammatory bowel disease: results of a British twin study." BMJ **312**(7023): 95-6.
Tian, Y., Y. Li, et al. (2010). "Differential effects of NOD2 polymorphisms on colorectal cancer risk: a meta-analysis." Int J Colorectal Dis **25**(2): 161-8.
Tominaga, E., H. Tsuda, et al. (2010). "Amplification of GNAS may be an independent, qualitative, and reproducible biomarker to predict progression-free survival in epithelial ovarian cancer." Gynecol Oncol **118**(2): 160-6.
Tomlinson, I. P., L. G. Carvajal-Carmona, et al. (2011). "Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer." PLoS Genet **7**(6): e1002105.
Tompkins, W. A., A. M. Watrach, et al. (1974). "Cultural and antigenic properties of newly established cell strains derived from adenocarcinomas of the human colon and rectum." J Natl Cancer Inst **52**(4): 1101-10.
Topilko, P., S. Schneider-Maunoury, et al. (1998). "Multiple pituitary and ovarian defects in Krox-24 (NGFI-A, Egr-1)-targeted mice." Mol Endocrinol **12**(1): 107-22.

Toyota, M., N. Ahuja, et al. (1999). "CpG island methylator phenotype in colorectal cancer." Proc Natl Acad Sci U S A **96**(15): 8681-6.
Toyota, M., H. Suzuki, et al. (2008). "Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer." Cancer Res **68**(11): 4123-32.
Travassos, L. H., L. A. Carneiro, et al. (2010). "Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry." Nat Immunol **11**(1): 55-62.
Trompeter, H. I., G. Blankenburg, et al. (1996). "Variable nuclear cytoplasmic distribution of the 11.5-kDa zinc-binding protein (parathymosin-alpha) and identification of a bipartite nuclear localization signal." J Biol Chem **271**(2): 1187-93.
Tur, G., E. I. Georgieva, et al. (2010). "Factor binding and chromatin modification in the promoter of murine Egr1 gene upon induction." Cell Mol Life Sci **67**(23): 4065-77.
UK National Statistics. (2011). "Colorectal (bowel) cancer." Retrieved 14/06/2011, 2011, from http://www.statistics.gov.uk/cci/nugget.asp?.id=2162.
van der Heide, F., A. Dijkstra, et al. (2009). "Effects of active and passive smoking on disease course of Crohn's disease and ulcerative colitis." Inflamm Bowel Dis **15**(8): 1199-207.
van Duijnhoven, F. J., H. B. Bueno-De-Mesquita, et al. (2009). "Fruit, vegetables, and colorectal cancer risk: the European Prospective Investigation into Cancer and Nutrition." Am J Clin Nutr **89**(5): 1441-52.
Vasen, H. F., P. Watson, et al. (1999). "New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC." Gastroenterology **116**(6): 1453-6.
Vermeire, S. and P. Rutgeerts (2005). "Current status of genetics research in inflammatory bowel disease." Genes Immun **6**(8): 637-45.
Virolle, T., E. D. Adamson, et al. (2001). "The Egr-1 transcription factor directly activates PTEN during irradiation-induced signalling." Nat Cell Biol **3**(12): 1124-8.
Vogelstein, B., E. R. Fearon, et al. (1988). "Genetic alterations during colorectal-tumor development." N Engl J Med **319**(9): 525-32.
Wagner, M., K. Schmelz, et al. (2008). "Transcriptional regulation of human survivin by early growth response (Egr)-1 transcription factor." Int J Cancer **122**(6): 1278-87.
Walther, A., R. Houlston, et al. (2008). "Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis." Gut **57**(7): 941-50.
Wan, P. T., M. J. Garnett, et al. (2004). "Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF." Cell **116**(6): 855-67.
Wang, G., X. Gao, et al. (2010). "Nucleophosmin/B23 inhibits Eg5-mediated microtubule depolymerization by inactivating its ATPase activity." J Biol Chem **285**(25): 19060-7.
Wang, W., H. S. Cho, et al. (2002). "Structural characterization of the reaction pathway in phosphoserine phosphatase: crystallographic "snapshots" of intermediate states." J Mol Biol **319**(2): 421-31.

Weisenberger, D. J., K. D. Siegmund, et al. (2006). "CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer." Nat Genet **38**(7): 787-93.
Weisz, L., A. Zalcenstein, et al. (2004). "Transactivation of the EGR1 gene contributes to mutant p53 gain of function." Cancer Res **64**(22): 8318-27.
Whitmarsh, A. J., P. Shore, et al. (1995). "Integration of MAP kinase signal transduction pathways at the serum response element." Science **269**(5222): 403-7.
Wiedmann, B., H. Sakai, et al. (1994). "A protein complex required for signal-sequence-specific sorting and translocation." Nature **370**(6489): 434-40.
Wu, M. Y., M. H. Chen, et al. (2001). "Experimental and clinicopathologic study on the relationship between transcription factor Egr-1 and esophageal carcinoma." World J Gastroenterol **7**(4): 490-5.
Xiao, J., Z. Zhang, et al. (2009). "Nucleophosmin/B23 interacts with p21WAF1/CIP1 and contributes to its stability." Cell Cycle **8**(6): 889-95.
Xie, J. and S. H. Itzkowitz (2008). "Cancer in inflammatory bowel disease." World J Gastroenterol **14**(3): 378-89.
Xie, T., A. Plagge, et al. (2006). "The alternative stimulatory G protein alpha-subunit XLalphas is a critical regulator of energy and glucose metabolism and sympathetic nerve activity in adult mice." J Biol Chem **281**(28): 18989-99.
Xu, Z., R. Dziarski, et al. (2001). "Bacterial peptidoglycan-induced tnf-alpha transcription is mediated through the transcription factors Egr-1, Elk-1, and NF-kappaB." J Immunol **167**(12): 6975-82.
Yamaguchi, K., S. H. Lee, et al. (2006). "Activating transcription factor 3 and early growth response 1 are the novel targets of LY294002 in a phosphatidylinositol 3-kinase-independent pathway." Cancer Res **66**(4): 2376-84.
Yan, Y. X., H. Nakagawa, et al. (1997). "Transforming growth factor-alpha enhances cyclin D1 transcription through the binding of early growth response protein to a cis-regulatory element in the cyclin D1 promoter." J Biol Chem **272**(52): 33181-90.
Yao, J., N. Mackman, et al. (1997). "Lipopolysaccharide induction of the tumor necrosis factor-alpha promoter in human monocytic cells. Regulation by Egr-1, c-Jun, and NF-kappaB transcription factors." J Biol Chem **272**(28): 17795-801.
Yoshimatsu, K., D. Golijanin, et al. (2001). "Inducible microsomal prostaglandin E synthase is overexpressed in colorectal adenomas and cancer." Clin Cancer Res **7**(12): 3971-6.
Yu, J., V. Baron, et al. (2007). "A network of p73, p53 and Egr1 is required for efficient apoptosis in tumor cells." Cell Death Differ **14**(3): 436-46.
Yu, J., I. de Belle, et al. (2004). "Coactivating factors p300 and CBP are transcriptionally crossregulated by Egr1 in prostate cells, leading to divergent responses." Mol Cell **15**(1): 83-94.
Yu, J., S. S. Zhang, et al. (2009). "PTEN regulation by Akt-EGR1-ARF-PTEN axis." EMBO J **28**(1): 21-33.

Zagurovskaya, M., M. M. Shareef, et al. (2009). "EGR-1 forms a complex with YAP-1 and upregulates Bax expression in irradiated prostate carcinoma cells." Oncogene **28**(8): 1121-31.
Zakharov, V. V. and M. I. Mosevitsky (2010). "Oligomeric structure of brain abundant proteins GAP-43 and BASP1." J Struct Biol **170**(3): 470-83.
Zhang, L., W. Zhou, et al. (1997). "Gene expression profiles in normal and cancer cells." Science **276**(5316): 1268-72.
Zhang, W. and S. Chen (2001). "EGR-1, a UV-inducible gene in p53(-/-) mouse cells." Exp Cell Res **266**(1): 21-30.
Zheng, C., Z. Ren, et al. (2009). "E2F1 Induces tumor cell survival via nuclear factor-kappaB-dependent induction of EGR1 transcription in prostate cancer cells." Cancer Res **69**(6): 2324-31.
Zheng, L., J. Pu, et al. (2010). "Abnormal expression of early growth response 1 in gastric cancer: association with tumor invasion, metastasis and heparanase transcription." Pathol Int **60**(4): 268-77.